Historia de la Probabilidad y la Estadística (IV)

Historia de la Probabilidad y la Estadística (IV)

Jesús Basulto Santos
Juan José García del Hoyo
(Eds.)

Revisión Técnica
María Dolores Pérez Hidalgo

Universidad de Huelva

A.H.E.P.E.

COLLECTANEA
131

2009
©
Servicio de Publicaciones
Universidad de Huelva
©
Jesús Basulto Santos
Juan José García del Hoyo
(Eds.)

Diseño de la cubierta y maquetación
Angel Gómez Rodríguez

Tipografía
Textos realizados en tipo Garamond de cuerpo 10, notas en Garamond
de cuerpo 8/auto y cabeceras en versalitas de cuerpo 10.

Papel
Offset industrial ahuesado de 80 g/m^2
Papel ecológico, exento de cloro

Encuadernación
Rústica, cosido con hilo vegetal

Printed in Spain. Impreso en España.

I.S.B.N.
978-84-96826-94-6

Depósito legal
H-41-2009

Imprime
Artes Gráficas Bonanza, S.L.

C.E.P.
Biblioteca Universitaria

Congreso Internacional de Historia de la Estadística y la Probabilidad (4º. 2007. Huelva y Sevilla)

Historia de la probabilidad y la estadística (IV) / edición de Jesús Basulto Santos, Juan José García del Hoyo. – Huelva : Universidad de Huelva, 2008

44 p.; 24 cm. – (Collectánea (Universidad de Huelva) ; 131)

"En este cuarto volumen se recopilan todas las ponencias presentadas al IV Congreso Internacional de Historia de la Estadística y de la Probabilidad, organizada por la Asociación de Historia de la Estadística y de la Probabilidad de España, y celebrado en Huelva y Sevilla durante los días 13 y 14 de septiembre de 2007". -- prólogo

ISBN 978-84-96826-94-6

1. Probabilidades – Congresos. 2. Estadística matemáticas – Congresos. I. Basulto Santos, Jesús. II. García del Hoyo, Juan José. III. Título. IV. Serie

519.2(063)

Prólogo

En este cuarto volumen se recopilan todas las ponencias presentadas al IV Congreso Internacional de Historia de la Estadística y de la Probabilidad, organizado por la Asociación de Historia de la Estadística y de la Probabilidad de España (AHEPE) y celebrado en Huelva y Sevilla durante los días 13 y 14 de septiembre de 2007, bajo el auspicio del Departamento de Economía General y Estadística de la Universidad de Huelva y del Departamento de Economía Aplicada I de la Universidad de Sevilla.

Es especialmente de agradecer la extraordinaria acogida que ambas universidades tuvieron hacia los participantes del Congreso. Unido a lo anterior, también es digno de mención no sólo la gran calidad de los trabajos realizados sino también su presentación y discusión en las que se puso de manifiesto un verdadero interés en los distintos tópicos de la Historia de la Estadística y de la Probabilidad, como viene siendo habitual desde el año 2001, periodo durante el cual se han celebrado cuatro congresos con una periodicidad bianual.

En este recorrido podríamos destacar distintos aspectos sobre el estudio de la Historia de la Estadística y de la Probabilidad en nuestro país:

- En primer lugar, la mayor parte de los grupos de investigación o investigadores individuales españoles dedicados o interesados en algún aspecto relacionado con la Historia de la Estadística y de la Probabilidad han tenido la oportunidad de intensificar y colaborar de un modo más eficaz y productivo en sus propias investigaciones e, incluso, en algunos casos, ha permitido tener conocimiento de la propia existencia de otros investigadores con temas de estudio afines.
- En segundo lugar, ha permitido estrechar e iniciar, en algunos casos, las relaciones con investigadores e instituciones extranjeras relacionados con dichos tópicos, lo que ha propiciado un continuo flujo científico enriquecedor en el ámbito del estudio de la Historia de la Estadística y de la Probabilidad.
- En tercer lugar, el devenir de estos cuatro congresos ha mostrado un creciente grado de madurez en los trabajos presentados en los mismos, como se puede apreciar en el cada vez más elevado nivel científico de los mismos.

En cuanto a los trabajos presentados lo primero que hay que decir es que superaron la treintena y versaron sobre muy variadas temáticas. La clasificación de los mismos que a continuación se detalla se ha realizado teniendo en cuenta algún nexo de unión con el epígrafe que los contiene, pero en ningún caso esto supone un compartimento estanco pues hay trabajos que comparten puntos en común con otro u otros de los epígrafes de esta clasificación.

La clasificación de los trabajos se puede realizar atendiendo a los siguientes epígrafes:

- La aplicación histórica de la probabilidad a cuestiones de carácter demográfico
- Trabajos de carácter biográfico o institucional
- Probabilidad y azar
- Tratados sobre Estadística y enseñanza de la Estadística
- Trabajos sobre la historia de diversos tópicos de la Estadística

A continuación se especificarán los títulos de los trabajos así como los nombres de sus autores dentro de cada uno de aquellos epígrafes.

La aplicación histórica de la probabilidad a cuestiones de carácter demográfico

- *Modéles pour la mesure du risque à l´àge classique* (Jean Marc Rohrbasser)
- *La Memoria de Daniel Bernoulli sobre la inoculación contra la viruela (1760): un problema de decisión bajo incertidumbre* (José Antonio Camuñez Ruiz, Jesús Basulto Santos y Francisco Javier Ortega Irizo)
- *Sur des tables de mortalité du XVIII siècle: écarts et similitudes* (Marc Barbut)
- *The aftermath of Abraham de Moivre´s Doctrine of Annuities on lives in 18th century Europe* (Ivo Schneider)
- *Reproducción y muerte de la población mexicana: cálculos estadísticos y preceptos higiénicos a fines del siglo diecinueve* (Laura Cházaro)
- *Orígenes de la bioestadística en España: estadísticas demográficas y sanitarias* (Antonio Franco Rodríguez-Lázaro y Mercedes Casas Guillén)
- *El estudio de la mortalidad en España a principios del siglo XX* (Sonia de Paz Cobo y Juan Manuel López Zafra)
- *Historia de las tablas de mortalidad españolas y su evolución* (Juan Escuder Bueno, Roberto Escuder Vallés y Ángel Vegas Montaner)

Trabajos de carácter biográfico o institucional

- *Maurice Halbwachs y la Estadística* (José María Arribas Macho)
- *Homenaje al profesor Dr. D. Francisco Azorín Poch* (Antonio Franco Rodríguez-Lázaro, Mª Carmen Escribano Ródenas y Andrés Gutiérrez Gómez)

- *El Instituto Nacional de Estadística de Lisboa. Un recorrido a través de su historia: orígenes, creación, funcionamiento y evolución* (Gabriela Fernández Barberis y Mª Escribano Ródenas)
- *Kolmogorov y sus aportaciones fundamentales* (Cristina Sánchez Figueroa, Pedro Cortiñas Vázquez e Iñigo Tejera Martín)
- *Kart Pearson, creador de la estadística matemática* (Miguel Ángel Gómez Villegas)
- *La disputa entre José de Alzate y el Virrey de la Nueva España, conde de Revillagigedo, por el censo de la ciudad de México de 1790* (Leticia Mayer Celis)
- *Aportación al conocimiento de la obra de Manuel Escudé Bartolí (1856-1930) impulsor de la estadística municipal en Barcelona* (Teresa Corbella Doménech y Manuel Escudé)
- *Olegario Fernández Baños, de la geometría a la econometría, sus aportaciones al análisis económico* (Salvador Almenar Palau)

Probabilidad y azar

- *El modelo del dado y su influencia sobre el desarrollo de la teoría de la probabilidad. Aciertos y fracasos* (Mª Sol de Mora Charles)
- *Una historia sobre las relaciones entre eficiencia y aleatoriedad* (José Javier Busto Guerrero y Jesús Muñoz San Miguel)
- *Una explicación de las regularidades detectadas por Pascal en su tabla de valores de las partidas (29 de agosto de 1654). Generalización a dos jugadores con diferentes habilidades* (Francisco Javier Ortega Irizo, Jesús Basulto Santos y José Antonio Camuñez Ruiz)
- *La resolución de Montmort (1710-1713) de los cinco problemas propuestos por Huygens en su tratado (1657)* (Mª Dolores Pérez Hidalgo y Jesús Basulto Santos)

Tratados sobre Estadística y enseñanza de la Estadística

- *J. Herrera Dávila y A. Alvear: Lecciones de Estadística (1829)* (Francisco Javier Martín-Pliego López y Jesús Santos del Cerro)
- *José María Ibáñez Ramos: primer Catedrático de Estadística* (Francisco Javier Martín-Pliego López)
- *Evolución de la Enseñanza de la Estadística en España a lo largo del siglo XIX* (Ana Isabel Busto Caballero y Mª Carmen Escribano Ródenas)
- *El libro de los dados de Alfonso X* (Jesús Basulto Santos, César Bordón Alba y José Antonio Camuñez Ruiz)

Trabajos sobre la historia de diversos tópicos de la Estadística

- *Lo que la astronomía regaló a la estadística* (Gabriel Ruiz Garzón)
- *Justificación de la regla de Borda: una revisión crítica* (Miguel Martínez Panero)

- *El desarrollo de las Estadísticas del sector pesquero durante los siglos XVIII y XIX* (Juan José García del Hoyo)
- *Aproximación a las poblaciones finitas en el contexto metodológico o histórico* (Santiago Murgui Izquierdo y Domingo Martín Martín)
- *La regresión por mínimos cuadrados parciales: orígenes y evolución* (Gregoria Mateos Aparicio-Morales y Antonio Jesús Caballero Domínguez)
- *Breve historia de la familia de clasificación Boosting* (Esteban Alfaro Cortés, Matías Gámez Martínez, Noelia García Rubio, José Luis Alfaro Navarro y José Mondéjar Jiménez)
- *Logit model: de Verhulst (1838) a McFadden (2001)* (Elena Martínez Rodríguez)
- *Investigaciones factoriales sobre la inteligencia. Técnicas en la década entre 1940 y 1950: aplicaciones a la selección de personal y a la orientación profesional* (Marcelo Pascual Faura)

Madrid, octubre de 2007

Dr. Francisco Javier Martín-Pliego López

Presidente de A.H.E.P.E.

Lista de Autores

Esteban Alfaro Cortés (457) Área de Estadística. F. CC. EE. y Empresariales de Albacete. Universidad de Castilla-La Mancha. E-mail: Esteban.Alfaro@clm.es.

José Luis Alfaro Navarro (457) Área de Estadística. F. CC. EE. y Empresariales de Albacete. Universidad de Castilla-La Mancha. E-mail: Jose.Luis.Alfaro@uclm.es.

José María Arriba Macho (31) Dpto. de Sociología I, UNED. E-mail: jarribas@poli.uned.es.

Marc Barbut (109) Centre d'Analyse et de Mathématiques Sociales, EHESS, París. E-mail: Marc.Barbut@ehess.fr.

Jesús Basulto Santos (47,185,357,383,407) Universidad de Sevilla. Dpto. de Economía Aplicada I. Avenida de Ramón y Cajal 1. E-mail: basulto@us.es.

César Bordon Alba (185) I.E.S. Punta del Verde. Sevilla. E-mail: cesarbordons@gmail.com.

Eric Brian (153) Directeur d'études à l'EHESS. E-mail: Eric.Brian@ens.fr.

Ana I. Busto Caballero (89) Universidad Complutense, Madrid.E-mail:aibusto@telefonica.net.

José Javier Busto Guerrero (337) Dpto. de Economía Aplicada I. Universidad de Sevilla. E-mail: jjbusto@us.es.

Antonio Jesús Caballero Domínguez (441) Jefe de Investigación de CFI Group. Madrid. E-mail: antoniocaballero@talk21.com.

José Antonio Camúñez (47,185,337,383) Universidad de Sevilla. Dpto. de Economía Aplicada I. E-mail: camunez@us.es.

Mercedes Casas Guillen (429) Universidad CEU San Pablo de Madrid. E-mail: mcasas@ceu.es.

Laura Cházaro García (395) Estudios Avanzados del Instututo Politécnico de Zacatenco. México. E-mail: chazaro@cinvestav.mx.

Teresa Corbella Doménech (421) Dpto. de Economía, URV. E-mail: tcd@fcee.urv.es.

Pedro Cortiñas Vázquez (257) Dpto. de Economía Aplicada y Estadística. UNED. E-mail: pcortinas@cee.uned.es.

Mª del Carmen Escribano Ródenas (77,99) Universidad CEU San Pablo, Madrid. E-mail: escrod@ceu.es.

Manuel Escudé Aixelà (421) Dpto. de Salud Pública, UB. E-mail: escude@ub.edu.

Juan Escuder Bueno (319) Universidad de Valencia. E-mail: Juan.escuder@uv.es.

Roberto Escuder Vallés (319) Universidad de Valencia. E-mail: roberto.escuder@uv.es.

Gabriela Fernández Barberis (99) Universidad CEU San Pablo, Madrid. E-mail: ferbar@ceu.es.

Antonio Franco Rodríguez-Lázaro (77,429) San Pablo CEU, Madrid. E-mail: fralaz@ceu.es.

Javier Gamero Rojas (481) Dpto. de Economía Aplicada I. Universidad de Sevilla. E-mail: jgam@us.es.

Matías Gámez Martínez (457) Área de Estadística. F. CC. EE. y Empresariales de Albacete. Universidad de Castilla-La Mancha. E-mail: Matías.Gámez @uclm.es.

Juan José García del Hoyo (265) Universidad de Huelva. E-mail: hoyo@uhu.es.

Noelia García Rubio (457) Área de Estadística. F. CC. EE. y Empresariales de Albacete. Universidad de Castilla-La Mancha. E-mail: Noelia.García @uclm.es.

Miguel Ángel Gómez Villegas (351) Universidad Complutense, Madrid. E-mail: ma_gv@mat.ucm.es.

Andrés Gutiérrez Gómez (77) Universidad CEU San Pablo, Madrid. E-mail: gutgom@ceu.es.

Juan Manuel López Zafra (311) Dpto. de Estadística e Investigación Operativa II. Fac. de CCEE, UCM. E-mail: juanma-lz@ccee.ucm.es.

Gregoria Mateos-Aparicio Morales (441) UCM. E-mail: goyi7@ccee.ucm.es.

Francisco Javier Martín-Pliego López (27 y 37) Universidad Rey Juan Carlos, Madrid. E-mail: jmartinp@mvaseguros.es.

Domingo Martín Martín (357) Dpto. de Economía Aplicada I. Universidad de Sevilla. E-mail: domartin@us.es.

Miguel Martínez Panero (241) Universidad de Valladolid. E-mail: mamecheve@unav.es.

Elena Martínez Rodríguez (449) Dpto. de Estadística e Investigación Operativa II, UCM. E-mail: emartinez@ccee.ucm.es.

Leticia Mayer Celis (375) Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM. E-mail: lmayer@servidor.unam.mx.

José Mondéjar Jiménez (457) Área de Estadística. UCLM. E-mail: Jose.Mondejar@uclm.es.

Marisol Mora Charles (175) UPV/EHU, San Sebastián. E-mail: demora@leibnizsociedad.org.

Jesús Muñoz San Miguel (337) Dpto. de Economía Aplicada I. Universidad de Sevilla. E-mail: jmiguel@us.es.

Santiago Murgui Izquierdo (357) Universidad de Valencia. E-mail: Santiago.Murgui@uv.es.

F. Javier Ortega Irizo (47, 383) Dpto. de Economía Aplicada I. Universidad de Sevilla. E-mail: fjortega@us.es.

Marcelo Pascual Faura (469) Dpto. de Empresa/Organización de Empresas, Universidad San Pablo-CEU, Madrid. E-mail: pascual.fcee@ceu.es.

Sonia de Paz Cobo (311) Dpto. de Estadística e Investigación Operativa II, UCM, Madrid. E-mail: depazcobo@yahoo.es.

María Dolores Pérez Hidalgo (407) Dpto. de Economía Aplicada I. Universidad de Sevilla. E-mail: mdperez@us.es.

Jean-Marc Rohrbasser (1) Institut National D'Études Démographies, Paris. E-mail: rohrbass@ined.fr

José Enrique Romero García (481) Dpto. de Economía Aplicada I. Universidad de Sevilla. E-mail: romerogje@us.es.

Gabriel Ruiz Garzón (221) Dpto. de Estadística e I.O. Universidad de Cádiz. E-mail: gabriel.ruiz@uca.es.

Cristina Sánchez Figueroa (257) Dpto. de Economía Aplicada y Estadística. UNED. E-mail: csanchez@cee.uned.es.

Jesús Santos del Cerro (27) Universidad Castilla – La Mancha. E-mail: jesus.scerro@uclm.es.

Ivo Schneider (231) Münchner Zertrum fúr Wissenschafts-und Technikgeschichte Deutches Museum. E-mail: vo.schneider@unibw.de.

Iñigo Tejera Martín (257) Dpto. de Economía Aplicada y Estadística. UNED. E-mail: itejera@cee.uned.es.

Ángel Vegas Montaner (319) Universidad de Alcalá. angel.vegas@uah.es.

CONTENIDO

Capítulo 1

Süssmilch et le risque: Femmes en couche et inoculation de la variole

Jean-Marc Rohrbasser
Institut National D'Etudes Démographies. Paris

Les raisonnements probabilistes touchant à des questions de population portent leur effort sur la compatibilité calculable d'une mesure du risque et du cours naturel ou providentiel des choses. A l'époque classique, on ne parle de « risque » que dans le domaine économique [Pradier, 2006]. Il serait donc plus juste, dans cette étude, d'évoquer une « mesure du sort », comme on disait alors, ou une « mesure des chances », à savoir, dans le cas présent, des chances de survivre ou de décéder. Pour des raisons pratiques, l'expression « mesure du risque » sera néanmoins maintenue [Bernstein, 1996]. Cette mesure soulève les difficultés liées à un ordre du monde humain avec ses valeurs ″flottantes″, ordre qu'il faut, tout présupposé qu'il soit, continuellement à nouveau démontrer en recherchant comment le suprême calculateur a combiné ces valeurs et ces nombres. Les techniques de prévision – notamment dans le domaine de la mortalité – opèrent à partir de cette problématique et dans cette visée : sur quoi peut-on fonder une science du contingent ?

Deux questions apparaissent centrales dans cette science du contingent naissante : celle de la mesure du risque de décès auquel sont soumises les femmes en couches, d'une part, les doutes portant sur la pratique de l'inoculation de la variole, d'autre part. Exposer ces deux cas d'école en prenant comme source les travaux de Johann Peter Süssmilch (1707-1767) donne l'opportunité d'une réflexion méthodologique sur ce que l'on appelle aujourd'hui la probabilité subjective ou psychologique. En évitant l'erreur classique d'une analyse de type rétrospectif qui introduirait dans les textes anciens des concepts qui n'étaient ni élaborés ni, lorsqu'ils existaient, les mêmes qu'aujourd'hui, il ne semble toutefois pas interdit, afin de mieux comprendre les pratiques et les notions du passé, de convoquer des considérations plus récentes qui, convenablement appliquées à la science de l'époque étudiée, peuvent se révéler éclairantes. Il faut prendre ces essais comme des expériences, des tests, des modèles dont la valeur heuristique n'est jamais évidente, mais probable.

I. Le cas des femmes en couches

Dans les deux éditions de son Ordre divin, Süssmilch détaille les causes qui font baisser la fécondité matrimoniale et envisage, parmi elles, « l'angoisse et la crainte du sexe féminin à cause des risques pour la vie lors de la naissance ». Le risque se calcule en prenant en compte le degré de l'incertitude relative au fonctionnement des lois naturelles, comme dans un jeu, un pari engagé contre la nature. Les mesures qu'il obtient, des probabilités issues d'analyses de fréquences, sont l'occasion, pour le théologien, d'exhorter à la confiance en la providence divine. En se posant la question – qu'il estime « nécessaire » — de savoir quel risque de décéder court une femme en couches, le pasteur de Berlin défend cette thèse que la crainte que ressent celle-ci est exagérée.

I. 1. Les faits

En 1741, Süssmilch considère la liste des décès de Londres de 1730 à 1739 et y dénombre les femmes « mortes soit d'un accouchement avant terme, soit dans les douleurs d'un accouchement à terme, soit encore après, pendant les couches [in denen Wochen], du fait d'autres accidents ». En vérifiant les calculs du pasteur à partir des bulletins de mortalité, on constate qu'il y a bien 2 498 femmes ainsi décédées. En rajoutant les fausses couches, on obtiendrait 2 534 décès en couches et un risque moyen de 1/71. Süssmilch est conscient du fait qu'il peut ainsi « juger […] du risque [Gefahr] que court une femme grosse et […] tant soit peu mesurer la grandeur de la crainte » en comparant « ce nombre avec le nombre de tous les enfants nés ». Or, pendant ces mêmes dix années, il y a eu 170 196 baptisés à Londres. Il ne s'agit bien que des baptisés, seul nombre figurant dans les bulletins. Le pasteur décide donc d'y ajouter « les mort-nés et les Abortus, ce qui [est] équitable puisque aucun d'entre eux n'est baptisé et que la somme précédente ne contient que les baptisés (christned) ». Il obtient 176 490 naissances en tout (cuadro 1).

Cuadro 1. Muertas de sobreparto y riesgo de fallecer en Londres del 1730 al 1739 (fuente : boletines de mortalidad)

Años	Muertas de sobreparto	Nacimientos	Riesgo de fallecer
1730	266	17 752	1/67
1731	251	18 503	1/74
1732	219	18 425	1/84
1733	292	18 121	1/62
1734	271	18 291	1/67
1735	192	17 463	1/91
1736	202	17 083	1/85
1737	284	17 401	1/61
1738	261	16 668	1/64
1739	260	16 783	1/65
Total	2 498	176 490	1/71

Süssmilch conclut que « les mortes en couches se rapportent à cette somme comme 1 à 69, ou encore que les mortes en couches représentent 1/70e de toutes les femmes en couches ; ce qui veut dire que, sur 70 femmes grosses, une y perd la vie ».

En 1761, Süssmilch utilise des chiffres provenant de Londres, d'Allemagne et d'Autriche (cuadros 2 y 3) :

Cuadro 2. Datos de Londres (en riesgos de fallecer)
(Datos corregidos después de comprobación)

Fuente	Riesgo
Graunt 1629-1636 et 1647-1658	1/51
Short 1629-1636	1/64
Short 1653-1660	1/36
Short 1734-1742 sans 1739*	1/70
Süssmilch 1728-1757*	1/75
* malpartos han sido añadidas.	

Cuadro 3. Datos alemanes y austriacos según Süssmilch y Baumann

Lugar y año	Riesgo
Berlin 1722-1724	1/92*
Berlin 1746	1/108*
Berlin 1757	1/102*
Berlin 1758-1774	1/89
Vienne 1738-1739	
Leipzig 1740-1749	1/65*
Leipzig 1759-1774	1/65
Gotha 1735-1751	1/70*
Gera 1740-1748	1/110*
Salzwdel y Arendsee 1766-1774	1/66
Lebus 1742-1775	1/73
140 pueblos de la Antigua Marcha 1766-1774	1/80
* datos de Süssmilch corregidos	

Aux relevés de Süssmilch s'ajoute la contribution de Nicolas Struyck (1687-1769) sur les « décès d'accouchées ». L'astronome hollandais soutient que « ce serait une méthode défectueuse que de déterminer le nombre [des décès d'accouchées] d'après les registres de Londres ou d'après d'autres registres imprimés ». Il utilise donc une statistique du village de Broek-in-Waterland en précisant le moment exact où le décès intervient après l'accouchement (cuadro 4).

Cuadro 4. Fallecimientos de paridas en Broek-in-Waterland
del principio de 1654 al 19 de octubre de 1742
[Struyck, 1753, V, 3 ; Struyck, 1912, p. 362]

Duración desde el parto	Numero de fallecimientos
24 horas	6
1 a 8 días	22
9 a 14 días	9
15 días a 3 semanas	7
22 días a 30 días	6
1 mes a 6 semanas	5
6 a 12 semanas	5
12 semanas a 3 meses	1

Il y a donc 61 décès et, indique Struyck, 1 923 accouchements. Le risque de décéder est donc ici de 1 décès pour 31,52 accouchements, soit 1/32. Struyck note également que, au cours des « 9 premiers jours après l'accouchement, il meurt autant d'accouchées que dans les 81 jours qui suivent » et, postulant que « de toutes les demoiselles qui se marient chaque année à Harlem et à Amsterdam la huitième ou la neuvième partie meurent pendant ou après l'accouchement », il recommande, comme à son accoutumée, d'essayer d'obtenir un résultat « sur de plus grands nombres » [Struyck, 1753-1912, V, 3, p. 362-363]. A contrario, Süssmilch note deux années – 1738 et 1739 — d'une liste viennoise qui montre un risque à peine croyable de 1 décès en couches sur 433 naissances. Le pasteur le transforme en 1/400, sans doute par souci de réalisme.

Baumann [Süssmilch, 1776, III, p. 107] corrige les calculs de son oncle en précisant que « cette comparaison des décès des femmes en couches et relevantes avec les naissances n'est pas entièrement exacte. Les enfants mort-nés auraient dû être additionnés aux baptisés, la mère et l'enfant succombant le plus souvent en même temps ». C'est bien ce que faisait Süssmilch en 1741 ; s'il ne l'a pas fait en 1761, ce n'est donc certainement pas méprise, mais bien plutôt manque des données nécessaires.

I. 2. L'interprétation des faits

En 1741, interprétant la proportion de 1/70, Süssmilch se place dans une perspective providentialiste : "¿no es estó poco si se consideran los peligros de todas las clases en los cuales se encuentra a una mujer durante el embarazo, el parto y el sobreparto? ¿No hay allí una razón papa agradecer a Dios de la protección particular que concede así al sexo femenino? ¿Y las mujeres tienen, por un temor impaciente, una razón de hacer la cosa más difícil de lo que es?" Ainsi, pour mieux défendre la thèse – évidemment nataliste — du moindre danger des couches, Süssmilch rappelle – en bon luthérien, lecteur précis de l'Ecriture — la véritable signification théologique et morale de la "malédiction" attachée à l'enfantement: "Es obviamente un castigo para el sexo femenino que debe tener los niños con dolor, pero este otro castigo, que las mujeres deben perder la vida en estos mismos dolores del parto, no figura en absoluto en la palabra divina. El dolor y la muerte son dos cosas bien diferentes" [Süssmilch, 1741, VII, 100, p. 302-303]. Toutefois, l'estimation du risque conduit à une prise de décision en situation d'incertitude : il convient d'aider la providence qui, en aucun cas incitation au fatalisme, n'exclut ni le libre-arbitre ni l'action humaine. Ainsi, la suite du raisonnement de Süssmilch est une condamnation de l'incurie — encore bien trop grande selon lui — des sages-femmes et des médecins incompétents. C'est pour mieux en revenir à la confiance qu'il faut accorder à la sagesse divine :

> *" […] una mujer, en su situación peligrosa, puede estar plenamente tranquila si, además de una mayor reflexión sobre lo que precede, está aún en condiciones de asegurarse de la divina providencia y la paternal asistencia de Dios. Esta última se muestra manifiestamente si se considera el caso en su conjunto, puesto que parece que de otro modo muchas más mujeres deberían fallecer en esta situación peligrosa. Quién puede, y particularmente en esto, proponerse esta asistencia es pues bienaventurado. Hay allí, para el sexo femenino, un motivo poco mediocre y determinante a la piedad ya que, así como este última es útil en todas las cosas, tiene del mismo modo, en particular en este peligro que la mayoría de las mujeres deben correr varias veces, una grande y poco común utilidad para rechazar todo el susto alarmante. " [Süssmilch, 1741, VII, 100, p. 303]*

Dès lors que Dieu a ordonné le monde et a prévu le décès accidentel en couches, deux stratégies sont possibles : soit laisser l'adversaire, c'est-à-dire les puissances trompeuses, agir pleinement et en toute irrationalité ; soit faire appel à une sorte de "probabilisme" raisonnable en mesurant le plus exactement possible la teneur de cet avenir dont la connaissance et le passage de la virtualité à la réalité n'appartiennent qu'à Dieu. Il s'agit bien de comprendre le danger pour moins le redouter. On passe ici d'une considération purement théologico-philosophique à une mesure, un nombre qui lui donnera un sens visant à déclencher un comportement — une décision — chez l'humain, demeurant, par la définition même de sa nature imparfaite, dans l'incertitude. Selon le pasteur, le risque couru par les femmes en couches, certes bien réel, n'en est pas moins fortement amplifié et de manière exagérée par l'imagination anticipatrice. Cette considération induit le raisonnement probabiliste classique inspiré du jeu de hasard qu'est la loterie: "¿ qué se pensaría de una persona que se viera a mitad muerto de alegría y pretendería ya construir castillos en España sobre la esperanza que tiene, en un juego, de sortear el lote ganando entre 70? Es sin embargo el mismo caso para las mujeres en el sobreparto que, sobre 70, una solamente debe morir, todos los 70 que se la representa todavía con mismo espanto" [Süssmilch, 1741, VII, 100, p. 303]. C'est articuler probabilité "fréquentiste" et probabilité "subjective", le nombre devenant un motif de croire.

Quelques considérations émises au XX^{e} siècle peuvent contribuer à approfondir la question telle qu'elle est posée par le pasteur de Berlin. Cette application de notions modernes à des données anciennes doit être effectuée avec prudence : il ne s'agit pas d'interpréter rétroactivement le passé avec du prétendu neuf, mais de voir en quoi une ancienne question peut nourrir nos propres doutes et nos propres interrogations.

Dans la conception statistique ou fréquentiste, la probabilité est définie comme la fréquence relative - ou la limite de fréquence relative - d'un type donné d'événements dans une série plus étendue du même type. La probabilité subjective ou psychologique entre en jeu chaque fois qu'une personne, dans n'importe quelle situation d'incertitude, formule un jugement, un choix, une préférence, prend une décision ou exprime sa confiance dans l'issue d'un événement donné. Dans cette conception, l'évaluation classique de la probabilité ne jouit d'aucun privilège, elle n'est que la plus habituelle ; l'indépendance elle-même n'est pas une propriété objective qui conditionne les possibilités de son bien-fondé. Maurice Fréchet schématise – ce sont ses propres termes - cette situation en distinguant une notion objective et une notion subjective de la probabilité. Dans le premier sens, c'est " una idealización de la noción de frecuencia - como la recta euclidiana es una idealización del borde de una regla ", dans le second, " el número llamado probabilidad de un acontecimiento es sólo la indicación numérica, el grado de creencia sobre la realización de este acontecimiento". L'articulation entre probabilité objective et probabilité psychologique constitue en soi un problème. Même si les considérations sur la seconde peuvent parfois étayer l'idée que l'esprit fonctionne comme un calculateur inconscient, les tendances qui président aux opérations mentales dans le domaine de la probabilité psychologique sont bien caractérisées et ne peuvent être assimilées aux règles de la probabilité mathématique. Elles se refusent à la réduction à une formule unique ou simple. Dans certaines circonstances, elles révèlent des propriétés analogues à celles de la probabilité mathématique, dans d'autres, elles s'en écartent de façon significative. Dans l'étude de la mortalité des femmes en couches, les différentes notions de la probabilité existent et interfèrent. Ce débat, qui se poursuit encore, est clairement désigné par Fréchet lorsqu'il décèle "sería chocante admitir que juicios subjectivos espontáneos puedan conducir a una apreciación numérica de grados de creencia que obedecerían a las reglas rígidas del Cálculo de las Probabilidades" [Fréchet, 1955, p. 303].

I. 3. Les éléments psychologiques de la probabilité

La probabilité psychologique est-elle mesurable et, si oui, par quels moyens? Pour le mathématicien anglais Frank Plumpton Ramsey, " es posible que lo que determina cómo deberíamos actuar también nos determine, directamente o indirectamente, tener una opinión justa sobre la manera en la que deberíamos actuar, sin que seamos hasta concientes de eso" [Ramsey, 1926-1931, p. 169][1]. Cette détermination, suivant Ramsey, consiste dans le principe de maximisation de l'utilité espérée. Les individus font des choix rationnels, choisissant l'acte qui a pour eux la plus grande valeur, étant donné leurs désirs et leurs degrés de croyance: "supondremos […] que nuestro sujeto tiene ciertas creencia con todo; entonces él actuará de modo que que él crea para ser las consecuencias totales de su acción será lo mejor posible" [Ramsey, 1926-1931, p. 177][2]. Il s'agit alors de mesurer les degrés de croyance comme bases d'actions possibles.

Comment les croyances d'un individu se coordonnent-elles aux degrés subjectifs de probabilité ? Le psychologue anglais John Cohen, fournit quelques indications à ce sujet. La réponse donnée par des fillettes de dix ans à la question "que significa la frase: ¿es probable que llueva?" est extrêmement variée, allant de "estoy segura que lloverá" à "yo no sé si lloverá o no", en passant par "hay muchas probabilidades para que llueva", "él podría llover", "me pregunto si lloverá". La moitié environ des enfants attribue à la phrase la signification "él es más probable que lloverá que el contrario, 5% de las niñas incluyendo que las probabilidades de lluvia son iguales a las de ausencia de lluvia". Pour des adultes,

[1] « It is possible that what determines how we should act determines us also directly or indirectly to have a correct opinion as to how we should act, without its ever coming into consciousness. »

[2] « […] we shall suppose […] that our subject has certain beliefs about everything; then he will act so that what he believes to be the total consequences of his action will be the best possible. »

l'interprétation de l'énoncé "el juez dice que el reo es probablemente culpable" se répartit comme suit (cuadro 5) :

Cuadro 5. Porcentajes de interpretaciones subjetivas de la palabra "probablemente" en la frase: "el juez dice que el reo es probablemente culpable". (Individuos adultos). (Según [Cohen, 1963, p. 155]).

Interpretación	Porcentaje
Cierto de la culpabilidad	28
Casi cierto de la culpabilidad sin el ser completamente	14
Más cierto de la culpabilidad que de la no culpabilidad	45
Probabilidades iguales de culpabilidad y no culpabilidad	13

Face à ce flou du langage, une solution est de recourir à l'intention de parier. Si un individu est prêt à parier à 3 contre 1 qu'il pleuvra demain, il est toujours loisible de poser que le degré de croyance qu'il assigne à cette proposition est de 3/3+1 = 0.75. Cependant, il faut encore connaître la nature des désirs et des valeurs que cet individu est prêt à associer à certains résultats de ses actions, d'une part, et, d'autre part, le degré plus ou moins élevé d'aversion qu'il peut avoir pour la prise de risque. La nécessité de ces compléments d'information est bien mise en lumière par Maurice Allais. L'économiste français souligne en effet que, parmi les éléments psychologiques qui interviennent dans la prise de risque, quatre sont absolument essentiels dans tout choix aléatoire.

1. En toda elección aleatoria, un individuo tiene en cuenta el valor psicológico vinculado a una ganancia, y no su valor monetario; " resulta de eso que si la satisfacción marginal es decreciente y si esta disminución es bastante fuerte, […] una ganancia décupla de otro podrá hasta el punto de vista psicológica tener sólo un valor doble, o posiblemente inferior ".
2. Un individuo tiene en cuenta las probabilidades tal como se los imagina, y no las probabilidades tal como están efectivamente.
3. Por la consideración de las esperanzas matemáticas, no se pueden sino aproximar las probabilidades de los valores psicológicos.
4. El individuo, según su grado de prudencia o temeridad, tiene en cuenta valores psicológicos según la forma de su distribución de probabilidad y en particular de su dispersión; " podemos jugar al póquer con mucha mayor fuerza posible siempre que el placer de participar en una combinación donde existen desviaciones es bastante fuerte para compensar la pérdida probable " [Allais, 1953, p. 507-510].

Une expérience proposée au XX^{e} siècle par l'économiste américain Daniel Ellsberg illustre clairement l'importance de la "probabilité" psychologique. L'un des enseignements de cette expérience est qu'il est impossible d'inférer des probabilités cohérentes avec la théorie à partir des choix de l'individu, lesquels s'avèrent, de ce point de vue, inconsistants : la cause en est que la décision prise dépend d'une appréciation subjective de la situation et non d'un raisonnement probabiliste conforme à la théorie. Ellsberg montre de plus le rôle fondamental que joue, dans la décision de l'individu, ce qui distingue une situation d'incertitude à risque estimé d'une autre à risque inconnu, même lorsque l'estimation, dans celle-là, est elle-même entachée d'ambiguïté et d'incertitude.

I. 4. L'expérience d'Ellsberg

L'économiste français Claude Henry propose une présentation simplifiée de l'expérience d'Ellsberg. Dans cette présentation, " un experimentador pone una muestra de sujetos ante la siguiente situación: una urna contiene 90 bolas que se sabe que 30 son rojas y el 60 otras azules o amarillas " sin que la proporción de estas dos últimas categorías sea conocida. Cada sujeto sortea una bola y puede elegir las siguientes (cuadro 6):

Cuadro 6. Experimento de Ellsberg en la presentación de Henry [2005]

		Roja	Azul	Amarilla
Elección 1	R	100	0	0
O	Az	0	100	0
Elección 2	$R \cup Am$	100	0	100
O	$Az \cup Am$	0	100	100

- En la elección 1, la mayoría de los sujetos optan por R>Az.
- En la elección 2, la mayoría de los sujetos optan por $Az \cup Am > R \cup Am$.
-Además, la mayoría de los sujetos optan simultáneamente por R>Az y $Az \cup Am > R \cup Am$. R>Az no puede ser compatible sino con desigualdad Pr (R)>Pr (Az) puesto que la ganancia es la misma para R y Az. Así mismo $Az \cup Am > R \cup Am$ no puede ser compatible sino con Pr $(Az \cup Am)$>Pr $(R \cup Am)$ por la misma razón. Pero, los sorteos siendo independientes, Pr $(Az \cup Am)$ = Pr (Az) + Pr (Am) y Pr $(R \cup Am)$ = Pr (R) + Pr (Am). Sería necesario pues tener simultáneamente Pr (R)>Pr (Az) y Pr (Az) Pr (Am)> Pr (R) + Pr (Am), lo que es imposible.

¿Por qué entonces estas elecciones? En la primera alternativa, la elección R>Az garantiza una ganancia con una probabilidad de 30/90 = 1/3. Si el sujeto supone que Pr (Az) = Pr (Am) = 0,5, la probabilidad de ganancia de la elección Az>R es también (90-60)/90 = 1/3. Pero el sujeto no es en absoluto cierto que haya tantas bolas azules (30) que de bolas amarillas (30) en la urna, a saber que Pr (Az) = Pr (Am) = 0,5. Huye de esta incertidumbre prefiriendo a R. Ocurre la misma cosa para la segunda alternativa: al hacer las elecciones que hacen, los sujetos del experimento huyen de la ambigüedad vinculada a la ignorancia de la proporción de las bolas azules y bolas amarillas [Henry, 2005, p. 17-20].

Les conclusions sont évidemment identiques en procédant suivant Ellsberg [1961]. Le sujet est mis en présence de deux urnes contenant des boules rouges et noires, de l'une desquelles il devra tirer une boule au hasard. "Parier sur Rouge 1" signifie qu'il choisit de tirer dans l'urne 1 et qu'il reçoit 100 € s'il tire une boule rouge ("si Rouge 1 arrive") et 0 € s'il tire une boule noire ("si non-Rouge 1 arrive"). Le sujet dispose de l'information suivante : l'urne 1 contient 100 boules rouges et noires, mais dans une proportion qu'il ignore : elle peut contenir aussi bien 1 que 99 boules rouges. Dans l'urne 2, le sujet sait qu'il y a exactement 50 boules rouges et 50 noires. Lorsqu'il choisit l'urne d'où la boule sera tirée, les probabilités subjectives étant les mêmes pour les deux urnes, le raisonnement en probabilités impliquerait que le parieur opte indifféremment pour l'urne 1 ou pour l'urne 2. Or, on constate que la réalité ne se conforme pas à la théorie : la plupart des parieurs expriment une préférence marquée pour l'urne 2. On trouvera une discussion approfondie de ces conclusions dans le cinquième chapitre de Ekeland [1991].

I. 5. Risque, providence et foi

L'individu agit tout naturellement en considérant l'ignorance comme un facteur de risque supplémentaire : il est manifeste que, quel que soit le pari, l'individu pense avoir plus de chances de gagner – ou de ne pas perdre – avec l'urne dont la composition est, sinon connue, du moins pas entièrement inconnue, ou avec une proportion de boules ne souffrant pas d'ambiguïté. Une situation d'incertitude à risque estimé est moins "inquiétante" ou "rebutante" qu'une situation d'incertitude à risque absolument inconnu ou dans laquelle une ambiguïté demeure sur les conditions du pari. Or, n'importe quel individu se trouve fréquemment placé dans une situation de pari : "durante toda nuestra vida, en un cierto sentido, nos engalanámos. Cada vez que vamos a la estación, asumimos la previsión de que un tren efectivamente se irá, y si no tuviéramos un grado suficiente de creencia en este acontecimiento, negaríamos la apuesta y nos quedaríamos en nuestra casa", comme l'écrit Ramsey. Et, poursuivant dans une logique que Süssmilch n'aurait certes pas désavouée, le philosophe affirme que "las opciones que Dios nos ofrece siempre son acondicionadas por nuestro poder de adivinar si una proposición cierta es verdadera" [Ramsey, 1926-1931, p. 183][3].

Dans la perspective où il s'agit de « deviner » les options offertes par Dieu, on comprend sans doute mieux, et peut-être autrement que comme une naïve profession de foi, la position providentialiste du pasteur de Berlin. Sans le dire explicitement, ce dernier paraît peu convaincu par le raisonnement mettant en avant la mesure du risque. Afin de persuader les femmes de se marier et d'avoir des enfants, le pasteur préfère faire s'évanouir toute crainte en recommandant de s'en remettre avec une totale confiance à la bienveillance divine. Le fait de montrer, par de bien maigres statistiques, que le risque couru est, selon lui, relativement faible ne doit certes pas conduire à refuser de parier pour l'enfantement sans pour autant garantir que ce pari soit absolument gagnant. Il ne le serait en effet qu'au prix d'un acte de foi qui ressemble davantage à l'engagement pascalien qu'à l'attitude d'un flambeur de casino. Certes, la mesure du risque – de la « chance » - de décéder doit conduire les femmes à " agradecer a Dios de la protección particular que concede así al sexo femenino " et à rassurer les femmes grosses. C'est bien plutôt le rôle déformant joué par la subjectivité, et en particulier par l'imagination, qui induit au raisonnement probabiliste déjà mentionné ci-dessus : ¿"qué se pensaría de una persona que se habría a mitad muerto de alegría y ya pretendería construir castillos en España sobre la esperanza que tiene, en un juego, de sortear el lote ganando entre 70? ". La rhétorique süssmilchienne, assimilant ici le risque de décéder à un « lot gagnant », se place peut-être dans la perspective d'une " infinidad de vida infinitamente feliz de ganar " [Pascal, 1662-1992, 680, p. 1213]. En tout état de cause, et même si le nombre devient un motif de croire en la représentativité de l'échantillon comme en la bienveillante toute-puissance de Dieu, le choix subjectif ne dépend pas, en dernière instance, d'un raisonnement en probabilités.

Que font donc les individus qui agissent comme s'ils n'assignaient aucune probabilité, quantitative ou qualitative, à l'événement auquel ils réagissent ? Dès lors que leurs choix ne

[3] « […] all our lives we are in a sense betting. Whenever we go to the station we are betting that a train will really run, and if we had not a sufficient degree of belief in this we should decline the bet and stay at home. The options God gives us are always conditional on our guessing whether a certain proposition is true. »

paraissent ni indifférents ni aléatoires, c'est bien que d'autres critères de décision interviennent. Dans le cas des femmes en couches, chacune d'entre elles peut toujours assigner une probabilité à sa survie ou à un accident mortel, probabilité qui traduit l'appui relatif donné par l'information dont elle dispose, son expérience et son intuition. Cela implique qu'elle peut toujours assigner des probabilités relatives à des états de la nature. Mais comment agira-t-elle dans cette situation ? La réponse à cette question peut dépendre d'une autre espèce de jugement, sur la fiabilité, la crédibilité ou la pertinence de son information ; non pas sur le soutien relatif que cela peut offrir à une hypothèse plutôt qu'à une autre contraire, mais sur sa capacité à soutenir quelque hypothèse que ce soit.

Si toute l'information sur les événements dans une série de paris était de la forme d'une distribution d'échantillon, l'ambiguïté serait directement en relation inverse avec la taille de l'échantillon. Mais celle-ci n'est pas un indice universellement utile de celle-là. L'information sur beaucoup d'événements ne peut pas être adéquatement décrite en termes de distribution d'échantillon ; en outre, la taille de l'échantillon semble principalement en rapport avec la quantité d'information. L'ambiguïté peut être élevée – et la confiance donnée à une quelconque distribution de probabilité faible – même lorsque l'information est abondante, lorsqu'il s'agit de sa fiabilité et de sa pertinence, et particulièrement lorsque il y a des opinions et des témoignages contradictoires. Ayant exploité la connaissance, la conjecture, la rumeur, la supposition, l'avis pour parvenir au jugement final qu'un événement est plus probable qu'un autre ou qu'ils sont également probables, on peut encore prendre du recul et se demander quelle est la valeur de tout cela, ce que l'on sait en définitive sur le problème, dans quelle mesure on a une base solide pour prendre une décision appropriée et agir pertinemment. La réponse : « je n'en sais pas beaucoup et je ne peux m'appuyer là-dessus » est passablement courante, même lorsqu'elle correspond à des estimations très inégales de probabilités relatives. Si l'ignorance complète est rare, voire inexistante, une ignorance « relativement grande » ne l'est certainement pas.

Cependant, à propos de la mortalité des femmes en couches, Süssmilch retrouve les accents piétistes d'une foi qu'il n'a jamais abandonnée dans la bienveillance de la providence. Lorsque l'on étudie les sciences de la nature, ¿"quién puede sondar", s'écrie le pasteur de Berlin, "las profundidades de la sabiduría, y mostrar la causa que hace que una cosa está así y no de otro modo, y no habría podido ser de otro modo?" Personne n'assiste au "conseil de Dieu" et ¿"nuestro conocimiento filosófico en su infancia, el orgullo pueril del hombre no nos hacen vergüenza?" A quoi bon cette prétention, ¿"no pueden confiarnos tranquilamente a la sabiduría del Creador que tiene todo hace así bien, tan bonito, así inconcebiblemente bonito y perfecto y espléndido de modo que el conjunto, en la resolución de lo que sea sobre tierra y de la intención que preside a la muchedumbre humana, debió tener su razón buena y suficiente?" [Süssmilch, 1761-62, I, Introduction, 8, p. 33] Il n'est pas interdit de percevoir dans ces accents le ton des partisans du retour à une simple piété qui avaient été, à Halle, les instituteurs du futur pasteur.

II. Les hasards de la variole

L'inoculation de la variole est médicalisée en Europe à partir de 1721, mais elle fait depuis longtemps l'objet d'une pratique empirique en Asie. La variole ne frappant jamais deux fois le même individu, il s'agit d'inoculer quelques gouttes de pus variolique, ce qui provoque l'éclosion d'une variole le plus souvent bénigne. Cependant, à l'époque et contrairement à la vaccine antivariolique qui fit son apparition à la fin du XVIIIe siècle, l'opération demeure une affaire risquée et parfois mortelle dans une proportion de 1/50 à 1/300.

II. 1. Les faits

Des estimations en chiffres absolus donnent grosso modo 60 millions de décès par variole au XVIII^{e} siècle, ce qui ferait de cette maladie l'un des premiers facteurs de mortalité. La mortalité variolique est très variable selon les régions et les périodes (cuadro 7).

Cuadro 7. Fallecimientos por viruela

Lugar y año	Numero de fallecimientos por viruela	Media anual
Nuremberg 1750	1 600 (niños)	
Paris 1716	14 000	
Paris 1723	20 000	
Montpellier 1744-45	2 000	
Rome 1752	6 000	
Turin 1755	800	
Berlin 1767	1 077	
Naples 1768	6 000	
Londres		2 082
Paris		1 428
Berlin		394
Edimbourg		122

De même, le risque de décéder de variole varie de 1/4 à 1/18 (cuadro 8).

Cuadro 8. Riesgo de fallecer por viruela

Lugar y año	Grado de mortalidad por viruela	Riesgo de fallecer por viruela
Edimbourg 1758-1762	10,52	1/10
Edimbourg 1743-1763	10,4	1/10
Besançon 1776-1777 et 1779	27,3	1/4
Dublin 1661-1690	21	1/5
Dublin 1715-1746	18,5	1/5
Berlin 1757-1774	8,3	1/12
Londres 1629-1636	2,8	1/36
Londres 1650-1699	5,8	1/17
Londres 1700-1719	6,8	1/15
Londres 1720-1739	8	1/13
Londres 1740-1759	8,4	1/12
Londres 1760-1779	9,3	1/11
Leyde 1765-1775	13,3	1/8
Ginebra (1) 1630-1679	6,7	1/15
Ginebra 1680-1729	5,7	1/18
Ginebra 1730-1780	5,7	1/18*

* El riesgo medio es de 1 fallecimiento por viruela para 13 fallecimientos. Darmon [1986] da un riesgo medio de 1/10.
(1) Las cifras de Ginebra provienen de Perrenoud [1980].

Après que Lady Mary Wortley Montagu a fait inoculer son fils en 1717 et sa fille en 1721, plusieurs médecins anglais décident d'adopter cette pratique. Le débat porte aussitôt sur les avantages et les risques de celle-ci. Voltaire, La Condamine, D'Alembert, Bernoulli et Lambert y prennent notamment part. Dès 1722, le médecin anglais James Jurin (1684-1750) pose nettement le problème : l'existence de quelques décès post-inoculatoires condamne-t-elle irrémédiablement cette pratique ? Jurin fait état de 2 décès pour 182 inoculations pratiquées par différents médecins et pasteurs en Angleterre et de 5 décès pour 300 à Boston : les valeurs de ces proportions peuvent justifier un refus de l'inoculation. L'argument qu'opposent les défenseurs de l'inoculation se fonde sur l'intervention d'autres causes pouvant expliquer ces décès. Il s'agit bien sûr de comparer deux risques, celui de mourir de la variole naturelle et celui de mourir de l'inoculation: "nuestra segunda intención es formar una estimación del riesgo que todo el género humano, los unos que llevan otros, corre de morir de las viruelas naturales; y, comparándolo con el riesgo de la inoculación, el público puede juzgar si sí o no la práctica de la inoculación tiende hacia la preservación del género humano disminuyendo el peligro al cual sería de otro modo sometido" [Jurin, 1722, p. 215-216][4].

Après avoir construit une table qui regroupe la mortalité générale et celle qui se produit du fait de la variole, Jurin, se fondant sur les bulletins de mortalité de Londres durant 42 ans, de 1667 à 1686 et de 1701 à 1722, estime le risque moyen de décéder lié à la variole naturelle à 1 décès sur 14. En réalité, ce risque varie avec l'âge. Jurin cherche également à savoir combien parmi les 13 autres personnes décédant ont contracté la variole sans en mourir. La mortalité par variole serait ainsi — mais le médecin anglais considère lui-même la proportion comme douteuse — de 1 décès par variole sur 8 sur 9 décès, la mesure que Daniel Bernoulli adoptera pour effectuer ses calculs. A la suite de cette avancée, si plusieurs médecins optent très vite pour l'inoculation, d'autres s'y opposent, ainsi qu'un apothicaire et des théologiens comme le rappelle La Condamine [1754-1759, p. 620-629]. En France, Voltaire se fait l'écho de la controverse anglaise et se déclare résolument favorable à l'inoculation :

> *"Se dice suavemente en la Europa cristiana que los Ingleses son locos y furiosos, locos, porque dan la viruela a sus niños para impedirles tenerla, furiosos, porque comunican de alegría a estos niños una enfermedad cierta y terrible para impedir un mal dudoso; los Ingleses por su parte dicen que los otros Europeos son flojos y desvirtuados, son flojos en lo que temen hacer un poco de mal a sus niños, desvirtuados, en lo que los exponen a morir un día de la viruela." [Voltaire, 1734, XI, p. 92-93]*

Voltaire donne aussi des chiffres pour justifier cette «insertion»: "sobre cien personas en el mundo, sesenta por lo menos tienen la viruela, de estos sesenta veinte mueren de eso en los años más favorables, y veinte lo conservan para siempre de restos lastimosos: he aquí pues la quinta parte de los hombres que esta enfermedad mata o desfigura seguramente" [Voltaire, 1734, XI, p. 101-102]. Toutefois la controverse ne s'amorce en France qu'en 1754 avec La Condamine, partisan inconditionnel de l'inoculation. Il répond aux six principales « objections » élevées contre cette pratique. La cinquième objection et les suivantes méritent d'être examinées.

> "Quinta Objeción. Es usurpar los derechos de la Divinidad que de dar una enfermedad o emprender de retirar aquél que en el orden de la Providencia se destinaba naturalmente

---

4 « […] the second Part of our Design […] is to form an Estimate of the Hazard which all Mankind, one with another, are under of dying of the natural Small Pox, that, by comparing this with the Hazard of Inoculation, the Publick maybe enabled to form a Judgment, whether or no the Practice of Inoculation tends to the Preservation of Mankind, by lessening the Danger to which they are otherwise liable. »

allí. […] Sexta Objeción. No está permitido dar una enfermedad cruel y peligrosa a alguien que quizá nunca no la habría tenido. […] Continuación de la misma objeción. […] ¿Podrá nunca convencer a un padre blando de comunicar deliberadamente a sus hijo único, una enfermedad que puede darle la muerte? ¿Algún pequeño que sea el riesgo al cual lo expone por la inoculación, si allí sólo tenía el sobre 100, sobre 200, sobre 300, como se lo supone, a que esta operación era inevitable, se debe exponerlo voluntariamente de este riesgo? " [La Condamine, 1754-1759, p. 634-647]

Objections qui, comme dans le cas des femmes en couches, convoquent réflexions théologico-philosophiques et considérations probabilistes.

II. 2. L'ordre de la providence

En 1761, Süssmilch développe un propos sur l'inoculation seulement esquissé vingt ans auparavant. Après avoir donné la proportion de 80 décès pour 1 000 – 1 décès par variole pour 12 ou 13 décès (12,5) — dus à cette maladie, taux que modifiera Lambert, Süssmilch s'adresse directement aux autorités politiques :

¿"[Esta práctica] no es de la más alta importancia y digna de toda la atención del Estado? ¿Se debería aquí ahorrar los gastos para dar a conocer exactamente este gran remedio y para ponerlo en uso? Sólo posible con ayuda de establecimientos públicos, con ayuda de experimentos suficientes y constantes sobre los pobres, y con ayuda de recompensas ofrecidas a los padres que llevarán a sus niños en dichos establecimientos. No es posible suprimir de otro modo los prejuicios de los padres. " [Süssmilch, 1761, I, XIII, 267, p. 532-533]

En effet, ajoute le pasteur, soucieux, comme dans le cas des femmes en couches, du rôle capital joué par la probabilité psychologique, ni la lecture d'ouvrages documentés sur la question, ni les calculs de la statistique probabiliste, ne convaincront de l'utilité du traitement. C'est donc à l'Etat de le promouvoir puisque le souverain, comme veut le démontrer le chapitre, a le devoir de conserver ses sujets en vie. Jugeant peut-être qu'il n'a été que relativement peu convaincant, le pasteur de Berlin revient à la charge dans une section beaucoup plus longue figurant dans le chapitre de son ouvrage expressément consacré aux maladies [Süssmilch, 1761, II, XXIV, 528]. La discussion porte essentiellement sur l'argument selon lequel user de l'inoculation constituerait une atteinte à la providence. C'est une réponse très circonstanciée à la cinquième objection rapportée par La Condamine. Il s'agit de refuser le fatum mahometanum, c'est-à-dire, pour Süssmilch comme pour le savant français, respecter l'ordre de la Providence et non s'y opposer :

"Pero, se opone, sería allí un ataque a la providencia. El Turco razona también así: Dios permite la peste, nadie puede morir sin la voluntad de Dios, la muerte es una verdadero bien para uno que cree en el Islam o un musulmán, es decir, aquél que a la verdadera doctrina, hay y la práctica. No se debe por lo tanto oponerse a la peste sino, al contrario, dejarle libre curso ya que, de otro modo, él se afectaría a la previsión y a la voluntad de Dios. " [Süssmilch, 1761, II, XXIV, 528, p. 440-441]

Une attitude fataliste reviendrait à confondre hasard et providence. Or, dans l'esprit du pasteur, il y a opposition radicale entre ces deux notions: "en un asunto tan importante como la conservación de la vida de un niño, no hay que dejar hacer por si acaso sino marchar al contrario con un paso asegurado. Si se hizo lo que podía hacer, la conciencia puede estar tranquila. En un asunto incierto, debemos escoger bien el más seguro y sacar a su niño del peligro más grande " [Süssmilch, 1761-62, II, XXIV, 528, p. 444]. Le plaidoyer du pasteur convoque ensuite les arguments scolastiques, repris par Leibniz [1710-1969], distinguant volonté divine parfaite et volonté permissive. Il poursuit avec des arguments probabilistes qu'utiliseront et approfondiront sur le plan mathématique D'Alembert, Bernoulli et Lambert, et qui portent sur le risque encouru lors de l'inoculation. Süssmilch, en toute bonne conscience et pour défendre la providence, n'hésite pas à mêler facteurs d'ordres psychologique, éthique et probabiliste :

> *"Pero, dicho aún, si uno de mis niños, o mi único niño, muere, no sólo seré inconsolable, sino que también deberé hacerme los mayores reproches mi vida durante. Es la palabra de una ternura privada de razón. ¿Si él muere en efecto según el curso ordinario de la viruela, no deberás también hacerte reproches? ¿Qué? No sería necesario decir esto: si se había servido de este remedio, habría tenido 25, o incluso 50 veces más de esperanza de salvar a tu niño; su salvación ya habría pasada a ser de vez más probable; habría disminuido el riesgo 25, o incluso de 50 veces; habría hecho al menos todo lo que era posible de lo que el deber y la razón lo habrían mandado y habría podido basarse en el hecho de realizar tu deber. Pero como, por un temor y una ternura desrazonables, dejó el riesgo aumentar en 25 a 50 veces, la responsabilidad de la muerte vuelve a caer sobre ti porque no prefirió más cierto al dudoso. " [Süssmilch, 1761, II, XXIV, 528, p. 444-445].*

Aussi Süssmilch conclut-il en fervent partisan de l'opération : il faut inoculer, c'est le devoir de chaque père et de chaque citoyen.

II. 3. Le problème de Newcomb

En 1960, un spécialiste de physique théorique aux Laboratoires des radiations de l'Université de Californie à Livermore, William Newcomb, construit un problème qui pone en escena un Ser que tiene la facultad de prever las elecciones humanas con una precisión casi absoluta: anteriormente, ha predecido estas elecciones con exactitud y nunca no se ha equivocado hasta ahora [Nozick, 1970; Gardner, 1973 et 1974 ; Watzlawick, 1976]. El Ser muestra dos cajas y explica que la caja 1 contiene 5 000 €, mientras que la caja 2 puede contener 50 000 € o nada. La alternativa es la siguiente: o tomar lo que se encuentra en las dos cajas (A), o no tomar más que el "contenido" de caja 2 (B). El Ser organizó las cosas de la siguiente manera: si la elección es (A), el Ser, que lo previo, dejará la caja 2 vacía, y la ganancia es de 5 000 €. Si la elección es (B), el Ser, que lo previo también, coloque los 50 000 € en la caja 2 (cuadro 9). El proceso se desarrolla en este orden: 1) el Ser opera su previsión; 2) en función de esta previsión, pone o no pone los 50 000 € en la caja 2; 3) el Ser comunica las condiciones de la operación; 4) el jugador hace su elección.

Cuadro 9. Las elecciones y las ganancias en el problema de Newcomb

	Caja 1	Caja 2
Elección A (1 et 2)	5 000	0 (o 50 000 ?)
Elección B (2)	0	50 000

Le joueur sait qu'il peut avoir une confiance presque totale dans la prescience de l'Être. S'il décide de choisir (A), l'Être l'aura prévu presque à coup sûr et aura laissé vide la boite 2.

Mais si la décision est (B), l'Être y aura presque certainement mis les 50 000 €. Il n'y a donc pas à hésiter : il est raisonnable de choisir (B). Or, suivant la séquence des événements, l'Être prévoit, puis informe des conditions, enfin le joueur choisit. Ainsi, au moment du choix, les 50 000 € sont ou ne sont pas dans la boite 2. Par conséquent, si la décision est (A), le gain est de 55 000 € ou au moins des 5 000 € présents dans la boite 1. Dans les deux cas, le choix (A) permet de gagner 5 000 € de plus. Il n'y a donc pas à hésiter : il est raisonnable de choisir (A). Cependant, les partisans du choix (B) peuvent faire remarquer que le raisonnement en faveur de (A) ne tient pas. L'Être l'a en effet presque à coup sûr prévu, et laissera vide la boite 2. Le gain n'est donc que de 5 000 € au lieu des 50 000 € réservés au choix (B).

Ce à quoi les partisans du choix (A) peuvent très bien opposer que l'Être, au moment du choix, a achevé sa prévision, et que les 50 000 € sont ou ne sont pas dans la boite 2, quelle que soit la décision prise. L'argent a été présent ou non depuis une heure, un jour, une semaine, avant la prise de décision. Cette dernière ne fera certes pas disparaître les 50 000 € s'ils sont dans la boite 2, de même qu'elle ne les fera pas apparaître s'ils n'y sont pas. C'est commettre une erreur, disent les partisans du choix (A), de supposer une causalité a posteriori impliquant que la décision du joueur est susceptible de faire apparaître ou disparaître les 50 000 €. Que l'argent soit ou non dans la boite 2, il est dans les deux cas déraisonnable de choisir (B) puisque, si la boite 2 est pleine, pourquoi négliger les 5 000 € présents dans la boite 1 ? Et s'il appert que la boite 2 est vide, le joueur peut se réjouir de gagner au moins 5 000 €. Testé sur plusieurs personnes, le problème de Newcomb répartit à-peu-près en nombre égal les partisans du choix (A) et ceux du choix (B).

Si, comme il se doit, on prend en compte le « presque à coup sûr » affectant la prévision de l'Être, le problème invite à une discussion sur l'induction. L'Être, qui ne s'est pas trompé jusqu'ici, peut très bien se tromper à présent puisque sa faculté de prévoir ne correspond qu'à une probabilité, certes très élevée, mais qui n'est pas la certitude. Le philosophe Hume montre qu'une proposition comme " el sol no se levantará mañana " est entièrement intelligible et " no impliques más contradicción que la afirmación: se levantará ". Il est donc parfaitement vain de tenter d'en démontrer la fausseté. Ce qui pousse le philosophe écossais à se demander quelle peut bien être "la naturaleza de esta evidencia que nos asegura de la realidad de una existencia o de un hecho más allá del testimonio actual de los sentidos o de los informes de nuestra memoria " [Hume, 1983, IV, 1, p. 85-86]. L'induction est entièrement tributaire de l'expérience et c'est tout ce dont le joueur dispose quant à la capacité de prévision de l'Être.

Dans les termes du problème de Newcomb, los partidarios de la elección (A) son los adversarios de la inoculación: en el momento de la elección, el Ser terminó su previsión, el individuo morirá o no morirá, y la inoculación no cambiará en eso nada. Aunque el experimento pone de manifiesto que, en casi todos los casos, salva una vida, el Ser puede no obstante equivocarse, y el experimento puede ser mortal. En cambio, la elección (B) implica una presciencia y una infalibilidad total del Ser así como una predeterminación inscrita en la necesidad de las cosas. Los partidarios de la elección (B) son los partidarios de la inoculación. Il est utile d'inoculer puisque l'on peut avoir une confiance entière en la prescience et en la bienveillance de l'Être.

Comme dans le cas des femmes en couches, la probabilité objective et la probabilité psychologique interfèrent : quel degré de croyance doit-on attribuer à la prescience et à la bienveillance de l'Être, et ce degré est-il mesurable ? Pour les adeptes du choix (B), cette croyance est sans faille, appuyée sur une certitude subjective absolue : il faut inoculer, cette pratique étant d'ailleurs inscrite de toute éternité dans le cours providentiel des choses. Les adeptes du choix (A) raisonnent, eux, en fonction d'une nécessité hypothétique, celle des futurs contingents, en fonction d'un degré de croyance relatif à une probabilité. Pour eux, l'Être est faillible, il a pu se tromper dans sa prévision, même avec une très faible probabilité, et l'inoculation ne guérit que presque certainement au prix de rares mais intolérables "bavures".

Les adeptes du choix (A) adoptent une attitude "pélagienne". Selon Pélage en effet, un individu peut toujours choisir entre le bien et le mal, et, pour ce faire, il dispose librement de son corps et de ses membres, sa volonté n'étant pleinement libre qu'en demeurant capable de ce choix. Pélage insiste sur l'autonomie de la créature, chef-d'œuvre du Créateur, puisque ce dernier lui a donné la raison, donc la conscience de ses actes.

Les adeptes du choix (B) adoptent "l'attitude luthérienne". Si l'individu croit que son choix est déterminé par l'ensemble des causes passées, l'idée de libre arbitre n'est qu'illusion dans un monde déterministe. Avec les doctrines du serf arbitre et de la prédestination, Luther proclame en effet la souveraineté absolue de Dieu. En ce monde règne la décision divine, et il est inutile de murmurer contre sa justice. Il faut bien plutôt que l'individu place en elle toute sa confiance, en craignant Dieu et en espérant en son sort, celui que la raison suffisante a déterminé quant à sa vie et à sa mort. C'est l'ultime réponse de Süssmilch à l'inquiétude des femmes en couches, c'est bien sûr également la sienne lorsqu'il s'agit d'inoculer la variole.

En se montrant un franc partisan de l'inoculation et — si imaginer sa réaction au problème de Newcomb a un sens — en optant pour le choix (B), Süssmilch réagit en bon luthérien. Le pasteur croit en la nécessité d'inoculer, c'est la providence même qui l'enjoint. Comme dans le cas des femmes en couches, Süssmilch invoque la bonté du Créateur, et l'accord de la volonté de la créature avec celle de la divinité. Mais il s'agit également d'une décision prise en fonction de probabilités estimées et d'un risque calculé lorsque sont prises en compte d'autres objections mentionnées par La Condamine.

II. 4. La mesure du risque

Les sixième, septième et huitième objections se rapportent à la mesure des risques relatifs de mourir de la variole naturelle et de l'inoculation. Pour La Condamine, cette réponse ne peut être fondée que sur les mathématiques puisqu'il n'est « plus question ici de morale ni de théologie, c'est une affaire de calcul : gardons-nous de faire un cas de conscience d'un problème d'arithmétique » [La Condamine, 1754-1759, p. 649]. Daniel Bernoulli (1700-1782) se propose de résoudre ce problème, relançant ainsi une controverse portant à la fois sur le calcul et l'éthique. Le mathématicien suisse considère deux risques, celui de contracter la variole naturelle et celui, l'ayant contractée, d'en mourir. Faisant l'hypothèse – qu'il sait simplificatrice - que ces deux risques sont constants et égaux à environ 1/8, il démontre alors que si l'on suppose, par exemple, « une génération de 13 mille enfants, il est sûr que si on pouvait les affranchir de la petite vérole, on sauverait par ce moyen la vie à environ mille de ces enfants ». En outre, poursuit Bernoulli, l'inoculation apporte, un surplus d'années d'espérance de vie, que l'on peut toutefois juger relativement peu élevé puisqu'il se monte à « environ 2 ans » [Bernoulli, 1760-1982, p. 240 et 235].

Au cours de ses réflexions sur le calcul des probabilités, D'Alembert revient sans cesse sur la question de l'inoculation [Paty, 1988, p. 203-265]. Sans condamner cette pratique, il répond au Mémoire de Bernoulli en faisant porter sa critique sur l'estimation même du risque que fait courir une inoculation. En premier lieu, la méthode de comparaison directe des risques ne lui paraît pas légitime :

"[…] suponiendo […] que el número de los que fallecen de la pequeña vérole sea 40 o 50 veces mayor que el número de los que se mueren de la inoculación, ¿se siga que los dos riesgos están el uno con el otro en la misma relación? La naturaleza del uno y de otro es bien diferente. Por pequeño que se quiera suponer el riesgo de morir de la inoculación, el que se hace inocular se somete a este riesgo en el corto espacio de 15 días, en el de 1 mes a lo sumo: al contrario el riesgo de morir de la viruela natural se extiende sobre todo el tiempo de la vida […] Si se quiere hacer […] un paralelo exacto de los dos riesgos, es necesario que el tiempo sea igual […] " [D'Alembert, 1761, p. 27-28].

C'est en second lieu une question très subjective d'estimation de son propre intérêt : en effet, pour D'Alembert, le risque présent — celui de décéder du fait de l'inoculation — ne saurait être compensé par un gain d'espérance de vie dont on ne profiterait qu'à terme. Le mathématicien français rappelle que, « suivant les tables de mortalité connues », un homme de 30 ans peut encore espérer vivre 30 ans. Qu'il puisse espérer en vivre 4 de plus ne justifie pas, à ses yeux, le risque pris en se faisant inoculer :

"[Suponen] que sometiéndose [a la inoculación] su vida media será de 34 años, es decir, de 4 años más que si esperara la viruela. Supongo por fin, con el Sr. Bernoulli, que el riesgo de morir de la inoculación será de 1 sobre 200. A partir de eso, me parece que para apreciar la ventaja de la inoculación, es necesario comparar, no la vida media de 34 años a la vida media de 30,sino el riesgo de 1 sobre 200, al cual se expone de morir en 1 mes por la inoculación (y eso a la edad de 30 años, en la fuerza de la salud y juventud), a la ventaja alejada de vivir más 4 años de al cabo de 60 años, cuando se esté mucho menos en condiciones de gozar de la vida. " [D'Alembert, 1761, p. 33].

Puisque D'Alembert pose que ces années gagnées le sont en fin de vie, une actualisation s'impose et l'utilité actuelle de ce supplément de vie diminue. Cela revient à introduire ce que plus tard l'on appellera l'utilité espérée. Le mathématicien jette ici les bases d'une théorie – subjective - de la valeur, que les actuaires énoncent ainsi : « tout gain ou toute perte futurs ont une valeur moindre, au moment présent, qu'un gain ou une perte actuels ». D'Alembert élève une autre objection, en rien mathématique, mais reposant sur le risque encouru : même si l'on sauve nombre de vies en inoculant, il suffit d'un décès consécutif à l'opération pour ôter toute pertinence à celle-ci. C'est la question de la mesure appliquée à l'individu comme membre d'une collectivité, la question même posée, sur un plan plus théorique, par la pratique de la statistique et des sciences sociales. Ainsi Condorcet, dans l'éloge de Daniel Bernoulli qu'il prononce en 1782, met l'accent sur cet autre nœud de la controverse, l'irréductibilité de l'intérêt collectif à une somme d'intérêts individuels :

"En 1760, Sr. Bernoulli aplicó el cálculo de probabilidades a la inoculación; vive esta cuestión en hombre público, y no se puede negar que no haya establecido de una manera victoriosa, y por un análisis muy fino, las ventajas de esta operación para un Estado donde se adoptaría generalmente; pero no la previo relativamente a cada particular. Bajo esta opinión, la cuestión cambia: en efecto, si un gran número de hombres se hacen inocular en un día, importa poco al interés general que una pequeña parte de estos hombres corre el riesgo de perder la vida al cabo de algunos días, puesto que el Estado compra a este precio una clase de certeza de conservar mucho más tiempo los que escaparán a este ligero peligro. No puede decirse lo mismo de cada particular; se trata, para él, de comparar un riesgo muy pequeño, pero próximo y estrechado en un espacio de tiempo muy corto, de un riesgo mayor, pero distante y

extendido sobre toda la duración de la vida. Pero Sr. Bernoulli sólo había calculado los efectos de la inoculación como un republicano a los ojos de quien el Estado es todo, y para el que los hombres no son más que ciudadanos. " [Condorcet, 1782-1994, p. 200].

Ces sages propos succèdent à un règlement mathématique de la controverse, règlement auquel avait essentiellement contribué Johann Heinrich Lambert (1728-1777). En se fondant sur les travaux de Bernoulli, le savant alsacien considère la dimension mathématique du problème de l'inoculation. Il n'entre pas dans la dispute elle-même mais entend mieux apprécier les risques relatifs. Chaque éventualité est caractérisée par une valeur numérique qui, lorsqu'elle est connue, est empruntée à Süssmilch (cuadro 10).

Cuadro 10. Combinación de los riesgos de morbosidad
y mortalidad por viruela, según Lambert

Morbosidad Mortalidad	Tuvieron la viruela	No tuvieron la viruela
Se murieron de la viruela	80	-
Se murieron de otra causa	1000 – a – 80	a

Lambert montre que Süssmilch n'a pas indiqué le bon taux des enfants mourant de la variole. Les 80 décès par variole sur 1000 décès que postule le pasteur de Berlin sont « manifestement trop peu. En effet, sur les 1000 qui meurent annuellement, il y en a un très grand nombre qui n'avaient pas contracté la variole. Supposons que leur nombre soit = a ; ainsi, sur les 1000 qui meurent chaque année, il y en a 1000 – a – 80 qui meurent ayant réchappé à la variole et 80 qui en meurent. Donc 80 meurent sur les 1000 – a qui ont été attaqués par la variole ». Sur cette base, Lambert commente et discute le travail de Bernoulli. Ce dernier avait considéré deux risques : celui de contracter la variole naturelle – Lambert l'écrit 1/l - et celui, l'ayant contractée, d'en mourir - 1/m. On se rappelle que le mathématicien suisse faisait l'hypothèse, d'une part, de la constance de ces risques aux différents âges, d'autre part, du poids égal de ces deux risques en retenant la valeur de 8 pour m et l. Lambert rejette ces hypothèses :

"Los niños quienes se deja retozarse en libertad a partir de su tercer o cuarto año son por lo tanto más propensos a la viruela que los más jóvenes que se mece aún, que se lleva o a que se tiene la brida. Parece resultar de esta observación que, los años pasando, 1/l aumenta, al menos hasta una determinada edad. [...]. Va diferentemente del m. En efecto, 1/m representa la medida del riesgo que se corre de morir de la viruela, o también la medida de la mortalidad por la viruela. Pues m, al igual que l, se implica en función de las distintas clases de edades, puesto que la viruela es más peligrosa, o virulenta, o mortal un año que otra. " [Lambert, 1772-2006, 129, p. 115].

S'appuyant sur des données provenant de Winterthur, Lambert montre que la valeur de m varie avec l'âge. A partir de ces chiffres, il teste la vraisemblance d'une constance et d'une égalité entre m et l et remarque que le risque de contracter la variole dépend étroitement des précautions prises pour l'éviter et que celles-ci sont susceptibles de varier notablement :

> *"Si se pudiera abstraerse de estas circunstancias, para la inmensa mayoría exteriores, con las cuales también se relaciona la diversidad de los países y de las maneras de vivir, podríamos concluir con certeza que cuanto más las viruelas alcanzan a niños y más mueren de eso, pues que l está a m en una proporción constante. Pero no podemos fácilmente dejar a un lado las circunstancias mencionadas. Y así como éstas son muy variables, los valores de l y de m no pueden ser cerrados lo mismo en límites estrechos. " [Lambert, 1772-2006, 140, p. 125].*

Ainsi, Lambert estime que l'hypothèse de Bernoulli doit être rejetée au moins pour les dix premières années de la vie. Par exemple, entre la naissance et le premier anniversaire, le risque de contracter la variole est de 1/39 tandis que le risque d'en décéder est de 1/3. En revanche, vers la quatrième année de vie, l'hypothèse de Bernoulli paraît acceptable.

II. 5. L'interprétation de la mesure

Dans la controverse sur l'inoculation de la variole, il s'agit encore de prendre des décisions, individuelles ou collectives, mettant directement en jeu la vie humaine dans des situations d'incertitude. Il convient donc de s'attacher au mécanisme de la probabilité ici en jeu et, comme pour les femmes en couches, à l'approche psychologique du risque, en convoquant quelques réflexions postérieures à celles de Süssmilch.

En abordant les conditions qui autorisent l'inférence d'une prémisse A à une conclusion B dans la perspective d'une décision à prendre en situation d'incertitude et d'un choix raisonnable qui doit s'ensuivre, le philosophe américain Charles Sanders Peirce poursuit la discussion inaugurée par Hume [Landemore, 2004]. Il rappelle que "la idea de probabilidad esencialmente pertenece a una especie de inferencia indefinidamente repetida"[5]. Ce qui implique qu'une "inferencia única necesariamente debe ser verdadera o torcida (falsea) y no puede mostrar ningún efecto de probabilidad"[6]. C'est bien là le problème posé par l'inférence : "si je me fais inoculer, alors je m'assure davantage d'espérance de vie : je me ferai inoculer une seule et unique fois et le risque mortel que je cours ce faisant a lui aussi un caractère, pour ainsi dire, "définitif". Peut-on alors parler de probabilité ? Non, répond Peirce, dès lors que, "considerado un solo caso en él mismo, la probabilidad no puede tener ningún significado»[7]. Poursuivant son analyse, le philosophe ajoute alors des réflexions d'ordre pragmatique tout à fait applicables à la question de l'inoculation :

> *"Si un hombre tuviera que elegir entre sortear 1 carta de un paquete conteniendo 25 cartas rojas y una sola negra, o de un paquete conteniendo 25 cartas negras y una sola roja, y si el sorteo de una carta roja estaba destinada a concederle una felicidad eterna mientras que el de una carta negra lo dedicaría a una desdicha sempiterna, sería locura negar que debería preferir el paquete conteniendo la mayor proporción de cartas rojas, aunque, según la naturaleza del riesgo, el sorteo no podría repetirse." [Peirce 1878, p. 160]* [8]

---

[5] « The idea of probability essentially belongs to a kind of inference which is repeated indefinitely. »

[6] « […] individual inference must be either true or false, and can show no effect of probability ».

[7] « […] a single case considered in itself, probability can have no meaning ».

[8] « […] if a man had to choose between drawing a card from a pack containing twenty-five red cards and a black one, or from a pack containing twenty-five black cards and a red one, and if the drawing of a red card were

Une transposition dans les termes de la question de l'inoculation pourrait être la suivante :

> *"Volvamos de nuevo al padre que duda en hacer inocular sus hijos; es a él que dirijo la palabra. Es, se dice, de la vida de sus hijos y no quieren aventurar nada. Tendrían razón seguramente si la cosa dependiera ustedes; pero es necesario aventurar aquí a pesar suyo: es en vano que se defienden. Sólo tienen dos partidos que tomar: o de inocular sus hijos o de no inocularlo; aquí dos casualidades que deben correrse, que uno es inevitable. Al inocular sus hijos, contra 375 acontecimientos felices, es 1 a temer; al no inocularlo, hay más de 1 que apostar contra 7 que lo perderán. Este último riesgo es 50 veces mayor que otro: elija ahora y duda aún si se lo atreven." [La Condamine 1754, p. 654]*

Ou encore celle-ci:

> "La proporción de los que se mueren de la viruela natural a los que se murieron a causa de la inoculación son de 12 a 300, lo que representa de 1 a 25. El riesgo es pues 25 veces menor en los inoculados que en los ordinarios. La esperanza de escaparse de nuevo a la viruela es 25 veces mayor […] ¿Qué se siente en el deber pues elegir a un padre que gusta a su niño con una ternura razonable y que está preocupado de su conservación? […] Si se es sitiado por la viruela en un tiempo en que ésta es epidémica y perniciosa, 1 niño muere sobre 6. […] Ahora bien, la viruela siendo un mal por todas partes extendido, la cuestión es: *¿qué deben elegir la razón y el amor de los padres? ¿debe y se puede no elegir, entre dos males, el menor? ¿Lo más grande de todos los deberes no son conservar la vida de su niño y de disminuir tanto más que es posible, por cuidados convenientes, el peligro que corre su vida? " [Süssmilch 1761, II, XXIV, 528, p. 442-443].*

Peirce reconnaît volontiers que "no es fácil conciliar [esta preferencia] con […] el análisis de esta concepción del azar"[9]. Raison de plus pour poursuivre ladite analyse : si l'individu choisissait" ¿el paquete rojo pero sorteaba la mala carta, cuál consuelo tendría? "[10] En prenant les chiffres de Süssmilch, nettement plus défavorables à l'inoculation (cuadro 11), on peut faire l'hypothèse que, en ne considérant que le risque relatif (les enfants inoculés auraient – ou n'auraient pas – contracté la variole et, dans le premier cas, en seraient – ou pas – morts), 1 enfant sur 25 risque de mourir de la variole inoculée.

Cuadro 11. Riesgos comparados en La Condamine y en Süssmilch

	Fallecimiento por viruela natural	Fallecimiento por inoculación	Proporción de los riesgos
La Condamine-Lambert	1/7	1/375	1 contra 53
Süssmilch	1/12	1/300	1 contra 25

destined to transport him to eternal felicity, and that of a black one to consign him ton everlasting woe, it would be folly to deny that he ought to prefer the pack containing the larger proportion of red cards, although, from the nature of the risk, it could not be repeated. »

9 « It is not easy to reconcile this with our analysis of the conception of chance. »

10 « […] choose the red pack, and should draw the wrong card, what consolation would he have ? »

Quel réconfort peut attendre le père qui voit mourir son enfant inoculé tandis que 24 autres survivent? Que pourrait invoquer ce tireur de mauvaise carte? "Podría decir ", poursuit Peirce, "que actuó racionalmente y que esto muestra solamente que su racionalidad absolutamente no valía nada"[11]. Par ailleurs, "¿si debía escoger la buena carta, cómo podría mirar esto de otro modo que como un azar feliz?"[12] Le principe même de la prise de décision en situation d'incertitude est que l'avenir est à la fois "certain" et incertain lorsque les événements sont indépendants. Il est certain, dans les termes des "lois du hasard", que la probabilité d'obtenir "pile" après un nombre n d'occurrences ininterrompues de "face" est, quel que soit n, de ½ si la pièce n'est pas biaisée ; et il est incertain, dans la même proportion, que l'événement "pile" se réalise. Toutefois, de son point de vue subjectif, n'importe quel joueur, avec, par exemple, n supérieur à 10, ne pourra s'empêcher de penser, en toute violation des règles du calcul des probabilités, que "face" a davantage de chances de se produire que "pile" au coup suivant, donc est plus "probable". Ainsi, le joueur pariant sur "pile" au 11e coup aura le sentiment de prendre un risque beaucoup plus élevé qu'un autre parieur dans une suite de coups plus classique ; pure illusion dès lors que la probabilité a priori est strictement la même avant chaque jet, "ilusión que", écrit Louis Bachelier, "valora sobre todo para que se traslade involuntariamente al principio de los acontecimientos cuando no es alejada demasiado. Si, la víspera, el juego se hubiera acabado por una serie de diez rojos, los jugadores no tendrían la idea de comenzar por ponter sobre el negro, el origen sería demasiado alejado. " [Bachelier, 1914, XV, p. 151]. Toutefois - et ce n'est certes pas sans intérêt dans la perspective d'une probabilité psychologique -, dans la solution qu'il propose au paradoxe de Saint-Pétersbourg, D'Alembert distingue la probabilité « métaphysique » de ce qui est simplement possible, de la probabilité «physique», celle de ce qui peut effectivement se produire. Il suggère "de debilitar el peso de los acontecimientos raros más que proporcionalmente en total de lo posible […] así […] la probabilidad de obtener cara solamente en el nième golpe en lugar de ser medida por el cociente $1/2^n$, será dado por una fórmula donde interviene un coeficiente de amortiguación a" [D'Alembert, 1768, p. 74-76 ; Granger, 1976, p. 45]. Dans un mémoire ultérieur paru en 1780, le mathématicien propose de calculer la probabilité d'un coup à pile ou face en tenant compte d'un coefficient de dépendance par rapport aux résultats antérieurs [D'Alembert, 1780, p. 45-46].

Dans son propre cadre de pensée, Süssmilch présente une transposition des réflexions précédentes :

> *"Sería con todo posible, dicho, que mi niño pudiera estar entre los pocos electos que no contrajera la viruela. Pero este argumento no vale nada, porque no se puede nunca estar en la certitud ni el pasar a ser; por eso poco cierto que, en una lotería, de sortear el mayor entre 100 probabilidades. En la incertidumbre, el riesgo demostrado permanece en todo momento. Ahora bien, en cosas tan importantes como la conservación de la vida de un niño, no es necesario dejar hacer todo a la casualidad, sino ir al contrario de un paso garantizado." [Süssmilch, 1761-62, II, XXIV, 528, p. 444].*

Es necesario pues, desde un punto de vista pragmático, razonar contra la razón, aunque ninguna racionalidad, en el caso de una única apuesta, puede autorizar a basarse en la certeza que la inferencia "si A entonces B" es legítima. Para Süssmilch, como en el caso de las

11 « He might say that he had acted in accordance with reason, but that would only show that his reason was absolutely worthless. »

12 « […] if he should choose the right card, how could he regard it as anything but a happy accident ? »

mujeres en sobreparto, la confianza incondicional colocada en la rectitud de Dios ordena en definitiva.

Mais la discussion précédente a opéré un glissement conceptuel par rapport aux réflexions de Peirce : le 25^{e} enfant inoculé — ou le n^{ième} – n'est plus analysable en termes d'événement singulier puisqu'il est déjà fait appel à la répétition, donc au mécanisme des grands nombres, donc à un emploi légitime de la probabilité. Mais "aunque la probabilidad manifiesta probablemente su efecto en, digamos, 1 000 riesgos, y esto por una determinada proporción entre el número de los éxitos y fracasos, eso no vuelve de nuevo con todo que a decir que será a la larga ciertamente así. Pero el número de riesgos, el número de inferencias probables que un humano toma y efectúa en su vida entera, es un número finito, y no puede ser absolutamente cierto que el resultado medio concordará plenamente con las probabilidades." [Peirce 1878, p. 161][13]. Peirce soutient ici que la balance des succès et des échecs après prise de décision risquée ne correspond pas forcément, sur une durée de vie humaine, à une moyenne où les deux résultats se contrebalancent à peu près, comme le voudrait la théorie. Avec cette proposition, le philosophe américain rejoint certains aperçus que l'on doit à l'école écossaise, et bien sûr à Hume, au XVIII^{e} siècle. En effet, pour le parieur de Peirce, "tomando todos sus riesgos en bloque, no puede tener allí ninguna certeza que no corresponderán a un fracaso, y su caso no es diferente, a otra escala, del [del tirador de carta]. Es el resultado indudable de la teoría de las probabilidades que un jugador, si juega mucho tiempo, debe al fin ser arruinado" [Peirce, 1878-1955, p. 159-161] [14]. Cependant, n'en déplaise à la théorie, qui refuse de jouer ? Mais cette interrogation ne concerne plus l'individu face à l'éventualité de l'inoculation.

Quoiqu'il accorde une grande confiance à la stabilité des rapports statistiques – elle est garantie par la providence divine – et donc au raisonnement statistique en général – dont il contribue incontestablement à perfectionner les linéaments comme à fonder les justifications épistémologiques —, Süssmilch est bien davantage intéressé et concerné par l'individu faisant face à l'éventualité de l'inoculation et par la décision qu'il prendra, peut-être en fonction d'un raisonnement probabiliste, plus sûrement à l'issue d'un acte pragmatique de civisme et d'un acte de foi. En répondant à l'objection consistant à voir dans la lutte contre la variole une atteinte à la providence, donc à la volonté de Dieu, l'auteur de L'Ordre divin, en une référence très leibnizienne, décrit l'action de Dieu en ce monde :

> *"Cuánto es necesario ser escrupuloso en tales razonamientos y distinguir bien la voluntad perfecta y graciosa de Dios de su voluntad permisiva según la cual deja producirse similar mal, pero delante de que el hombre se tiene libre en el ejercicio de su deber de conservación de sí mismo, libre de resistir al mal así permitido y, confiando en la gracia divina, de utilizar con razón un remedio valioso y saludable contra el mal y la enfermedad" [Süssmilch, 1761-62, II, XXIV, 528, p. 440-441].*

---

13 « Although probability will probably manifest its effect in, say, a thousand risks, by a certain proportion between the numbers of successes and failures, yet this […] is only to say that it certainly will, at length, do so. Now the number of risks, the number of probable inferences, which a man draws in his whole life, is a finite one, and he cannot be absolutely certain that the mean result will accord with the probabilities at all. »

14 « Taking all the risks collectively, then, it cannot be certain that they will not fail, and this case does not differ, except in degree, from the one last supposed. It is an indubitable result of the theory of probabilities that every gambler, if he continues long enough, must ultimately be ruined. »

La volonté antécédente veut le bien, la volonté conséquente veut et actualise le meilleur possible : exemplaire illustration du maintien du libre arbitre humain, du moins de cette "illusion" puisqu'il s'agit d'un libre arbitre de type luthérien, prévu par la prescience de Dieu.

Süssmilch effectue une comparaison des deux risques mortels encourus, lors des couches et par l'inoculation. En effet, ceux qui redoutent de courir le risque de l'inoculation "no deberían casar a su hija, sobre todo si es única" puisque lorsqu'une femme se marie, elle est exposée au risque de mourir en couches. Quoique peu élevé, ce risque l'est encore beaucoup plus que pour l'inoculation de la variole, «plus de trois fois plus grand», précise le pasteur, alors qu'il faudrait même aller jusqu'à plus de 4 fois (300/70 ≈ 4 ¼). Est-ce une raison pour «empêcher les filles de se marier et le leur déconseiller»? Ce qui restaure la confiance est d'abord de penser "que la chica a la que se casa no será entre estas madres pobres cuya muerte es el precio de la vida de su niño ", ce qui correspond au tireur de carte peircien qui, dans un paquet comprenant 69 cartes rouges et une noire, ne doit surtout pas tirer cette dernière. Toutefois, en dépit de la non-valeur rationnelle de l'inférence, le joueur, dès lors qu'il connaît la composition du paquet de cartes, est dans une situation certes d'incertitude mais relativement confortable. Ce qui n'empêche pas, déclare Süssmilch, qu'il convient avant tout de « suivre l'ordre de Dieu dans la nature» et de songer que « le mariage est la vocation [Beruf] du sexe féminin ». On peut donc, dans le cas de l'inoculation avoir tout autant confiance, voire davantage, les raisons étant encore plus fortes puisqu'il faut "utilizar este remedio con la causa del bien común, con amor de la especie humana, con su propia patria de la que la población puede sólo ganar allí por todas partes, por fin por obediencia a la voluntad de Dios, la cual quiere que el hombre esté numerosos sobre la tierra" [Süssmilch, 1761-62, II, XXIV, 528, p. 445-446].

Une dernière comparaison de risques ne manque également pas d'intérêt si l'on se souvient que, très probablement, l'origine de l'idée d'assurance est à trouver dans les anciens contrats de navigation. Le péril en mer constitue en effet un risque majeur. C'est à ce dernier que Süssmilch confronte le risque attaché à l'inoculation de la variole. Pourrait-on, s'interroge le pasteur, faire du commerce et naviguer, «un père y destinerait-il son enfant s'il voulait regarder au risque?» Ce sont « les voyages au long cours vers les Indes orientales» qui apparaissent comme les plus dangereux, à la fois à cause de la traversée elle-même et des « nombreuses maladies spécifiques qui se déclarent lors d'un séjour prolongé sous l'équateur, par manque d'eau et pour beaucoup d'autres raisons». En outre, "la observación directa de un pobre naufragio durante el cual una embarcacíon zozobra, y a menudo solamente después de sufrimientos indescriptibles" ne devrait-elle pas faire réfléchir les plus courageux et faire cesser le commerce maritime? Süssmilch procède suivant sa méthode habituelle : empruntant ses chiffres à Struyck, il propose une estimation du risque et en tire les conclusions qui s'imposent (cuadro 12).

Cuadro 12. Medida de los distintos riesgos en Süssmilch

Viruela natural	Inoculación	Mujeres en sobreparto	Navegación
1/12	1/300	1/70	1/15

L'astronome hollandais ajoute que dans "un viaje total, ir y la vuelta, la quinta parte aproximadamente de la tripulación muere" mais qu'il ne faut pas croire "que así todo es dicho y que encontramos el verdadero informe […] estas cifras [siendo dados] que para hacer a otros autores examinar la cuestión más detalladamente" [Struyck, 1753-1912, p. 361-362]. Cependant, même si ces navigateurs se sacrifient, poursuit Süssmilch, "¿con el fin de proporcionar subsistencia y provecho a otros, pensamos decir que quienquiera, sea el que el moralista más riguroso, hubiera ascendido las protestas a este sujeto, mientras que no falten razones aparentes para hacerlo?". Dès lors que beaucoup de navires font chaque année ce commerce, "¿el riesgo no debería retener a los hombres, y las pérdidas humanas no deberían ofrecer bastantes motivos para repudiar totalmente tales expediciones?" Il n'en est rien tandis que, "con respecto al más inofensivo, del más seguro remedios, dejamos sus aprensiones ir tan lejos como tal la inoculación se ve prohibida como un entredicho medio, peligroso y que no se podría usar de eso con buena conciencia". Süssmilch termine par une comparaison avec la mortalité due aux guerres – qu'il ne chiffre pas – dont il tire les habituelles réflexions morales et civiques sur l'inoculation.

A l'issue de ces essais, non seulement de mesure du risque mais d'estimation subjective de sa valeur du point de vue de la psychologie de l'individu qui y est soumis, les questions afférentes à toute réflexion sur la probabilité psychologique demeurent posées : quel est son statut lorsque son estimation est confrontée à la prise de risque et de décision ? Quels résultats obtiendrait-on dans le cas de phénomènes naturels comme les catastrophes, dont on a pu récemment écrire qu'elles sont, dans une certaine mesure, elles aussi sous la responsabilité du comportement humain? Comment, en partant des calculs, des réflexions, des comparaisons qui viennent d'être rappelées, déterminer le seuil à partir duquel on peut ou non courir un risque, que l'on agisse en "décideur rationnel" ou par une forme d'instinct ou de subjectivité qui se rapproche davantage de la psychologie du parieur, du joueur? Qu'est-ce qu'un "décideur rationnel"?

Bibliographiques

ALLAIS MAURICE, «Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'Ecole américaine». Econometrica, Vol. 21, n° 4, octobre 1953, p. 503-546.

BACHELIER LOUIS, *Le Jeu, la Chance et le Hasard*. Paris, 1914.

BERNOULLI DANIEL, «Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir», Histoire de l'Académie Royale des Sciences, p. 1-45, III. 2, Paris, 1760, dans Die Werke von Daniel Bernoulli, vol. 2. Bâle, Boston Stuttgart, 1982.

BERNSTEIN PETER L., *Against the gods : The remarkable story of risk. New-York, 1996.*

COHEN JOHN, *Hasard, adresse et chance. La psychologie du pari et du jeu* [1960]. Traduction Elisabeth Grin. Paris, 1963.

CONDORCET JEAN-ANTOINE-NICOLAS DE CARITAT MARQUIS DE, «Eloge de Daniel Bernoulli» [1782] dans Condorcet, Arithmétique politique. Textes rares ou inédits (1767-1789), Paris, 1994, p. 200.

D'ALEMBERT JEAN LE ROND, *Opuscules mathématiques ou Mémoires sur différents sujets de Géométrie, de Méchanique, d'Optique, d'Astronomie etc*. Paris, 1761, « Onzième Mémoire : Sur l'application du calcul des Probabilités à l'inoculation de la petite Vérole », Tome II, p. 26-47.

D'ALEMBERT JEAN LE ROND, *Opuscules mathématiques ou Mémoires sur différents sujets de Géométrie, de Méchanique, d'Optique, d'Astronomie etc*. Paris, 1768, « Vingt-troisième

Mémoire : Extrait de plusieurs Lettres de l'Auteur sur différents sujets, écrites dans le courant de l'année 1767 ; V. Sur le calcul des probabilités », Tome IV, p. 73-79.

D'ALEMBERT JEAN LE ROND, *Opuscules mathématiques ou Mémoires sur différents sujets de Géométrie, de Méchanique, d'Optique, d'Astronomie etc*. Paris, 1780, « Cinquante-deuxième Mémoire : II. Sur le calcul des Probabilités ». Tome VII, p. 39-60.

DARMON PIERRE, *La longue traque de la variole. Les pionniers de la médecine préventive.* Paris, 1986.

EKELAND IVAR, *Au hasard. La chance, la science et le monde*. Paris, 1991.

ELLSBERG DANIEL, «Risk, Ambiguity, and the Savage Axioms». The Quarterly Journal of Economics, 75, n°4, 1961, p. 643-669.

FRECHET MAURICE, «Sur l'importance en économétrie de la distinction entre les probabilités rationnelles et irrationnelles» Econometrica, Vol. 23, n° 3, juillet 1955, p. 303-306.

GARDNER MARTIN «Reflections on Newcomb's Problem: A Prediction and Free-will Dilemma», Scientific American, n° 230, mars 1974, p. 102-108.

GARDNER MARTIN, «Free Will Revisited, with a Mind-bending Prediction Paradox by William Newcomb», Scientific American, n° 229, juillet 1973, p. 104-109.

GRANGER GILLES-GASTON, «Le paradoxe de Saint-Pétersbourg ». Le jeu au XVIIIe siècle. Actes du Colloque d'Aix-en-Provence, 1971. Aix-en-Provence, 1976, p. 43-48.

HENRY CLAUDE, «Du risqué à l'incertitude dans les modèles de décisions». Chaire de développement durable, EDF-Ecole Polytechnique, Cahier n° 2005-007. Avril 2005.

HUME DAVID, *Enquête sur l'entendement humain* [1748], traduction A. Leroy, Paris, 1983.

JURIN JAMES, « A letter to the Learned Dr. Caleb Cotesworth, F.R.S. of the College of Physicians, London, and Physician to St. Thomas's Hospital; containing a Comparison between the Danger of the Natural Small Pox, and of that given by Inoculation…», Philosophical Transactions, XXXII (1722-1723), p. 213-227.

LA CONDAMINE CHARLES-MARIE, « Mémoire sur l'inoculation de la petite vérole », Histoire de l'Académie royale des sciences. [Année 1754. Avec les Mémoires de Mathématique et de Physique pour la même année, tirés des registres de cette Académie]. Paris, 1759, p. 615-670.

LAMBERT JOHANN HEINRICH, *Beyträge zum Gebrauche der Mathematik und deren Anwendung,* Berlin, 1765-1772, Vol. III (1772), Section IX, pp. 476-599, Anmerkungen über die Sterblichkeit, Todtenlisten, Geburten und Ehen. J. H. Lambert, Contributions mathématiques à l'étude de la mortalité et de la nuptialmité (1765 et 1772). Edition critique, bilingue par Jean-Marc Rohrbasser et Jacques Véron. Suivi de « Les équations de Lambert » par Marc Barbut. Paris, 2006.

LANDEMORE HELENE, *Hume. Probabilité et choix raisonnable.* Paris, 2004.

LEIBNIZ GOTTFRIED WILHELM, *Essais de Théodicée sur la bonté de Dieu, la liberté de l'homme et l'origine du mal* [1710]. Ed. Jacques Brunschvig, Paris, 1969.

NOZICK ROBERT, « Newcomb's Problem and the Two Principles of Choice », in Essays in Honor of Carl G. Hempel, Dordrecht, 1970, p. 114-146.

PASCAL BLAISE, *Pensées* [1662]. Ed. Gérard Ferreyrolles et Philippe Sellier, Paris 1992.

PATY MICHEL, « d'Alembert et les probabilités ». Sciences à l'époque de la Révolution française (sous la dir. de R. Rashed). Paris, 1988, p. 203-265.

PEIRCE CHARLES SANDERS, « On the doctrine of chances, with later reflections » [1878]. Philosophical writings of Peirce. Ed. Justus Buchler, New York, 1955, p. 157-173.

PERRENOUD ALFRED, *La population de Genève du XVI^e au début du XIX^e siècle. Etude de démographie*. Genève, 1980.

PRADIER PIERRE-CHARLES, *La notion de risque en économie*. Paris, 2006.

RAMSEY FRANK PLUMPTON, «Truth and Probability» [1926]. The Foundations of Mathematics and other Logical Essays. Ed. R. B. Braithwaite, London, 1931, p. 156-198.

SAVAGE LEONARD JIMMIE, *The Foundations of Statistics*. New York, 1954.

STRUYCK NICOLAS, *Vervolg van de beschrijving der staartsterren, en nader ontdekkingen omtrent den staat van het menchelijk geslagt, uit ondervindingen opgemaakt, benevens eenige stermerkingen*, Amsterdam, 1753. Les Oeuvres de Nicolas Struyck (1687-1769) qui se rapportent au calcul des chances, à la statistique générale, à la statistique des décès et aux rentes viagères, tirées des Œuvres complètes, et traduites du hollandais par J. A. Vollgraff, [offertes aux membres du Septième Congrès international des actuaires, réunis à Amsterdam en septembre 1912 par la Société Générale Néérlandaise d'Assurances sur la vie et de Rentes viagères, établie à Amsterdam, Damrak, 74.] Amsterdam, 1912.

SÜSSMILCH JOHANN PETER, *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts, aus der Geburt, Tod, und Fortpflanzung desselben erwiesen* von Johann Peter Süßmilch..., Berlin, 1741. Traduction française de J.-M. Rohrbasser, L'Ordre divin dans les changements de l'espèce humaine, démontré par la naissance, la mort et la propagation de celle-ci. Paris, 1998.

SÜSSMILCH JOHANN PETER, *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts, aus der Geburt, dem Tode und der Fortpflanzung erwiesen von Johann Peter Süßmilch..., Zwote und ganz umgearbeitete Ausgabe*, Berlin, 1761-62.

SÜSSMILCH JOHANN PETER, *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts, aus der Geburt, dem Tode und der Fortpflanzung desselben erwiesen...*, Berlin, 1776. [Avec une Troisième partie qui contient des observations et des additions aux deux premières parties, et un traité sur les sociétés d'assistance aux veuves. Publiée par C. J. Baumann, pasteur à Lebus].

VOLTAIRE FRANÇOIS-MARIE AROUET, dit, «Onzième Lettre sur l'insertion de la petite vérole ». Lettres philosophiques. Amsterdam, 1734.

WATZLAWICK PAUL, *How real is real?* Communication, Disinformation, Confusion, New York, Toronto, 1976.

Capítulo 2

J. Herrera Dávila y A. Alvear: Lecciones de estadística(1829)

FRANCISCO JAVIER MARTÍN–PLIEGO LÓPEZ
Universidad Rey Juan Carlos
JESÚS SANTOS DEL CERRO
Universidad de Castilla La Mancha

Introducción

Se considera el tratado de José María Ibáñez, primer catedrático de Estadística en España en el seno de la Sociedad Matritense de Amigos del País, titulado *Tratado Elemental de Estadística* (Tomos I y II) publicados en 1844 y 1845 como el primer tratado sistemático de la nueva ciencia de la Estadística en nuestro país. Es una obra extensa en la que se trata una amplia temática y en la que se recogen numerosas referencias de otras obras sobre Estadística como las de Peuchet, Gioja y Dufau. No obstante, siendo este el primer referente en castellano de las obras escritas sobre la ciencia emergente de la Estadística, las *Lecciones de Estadística* de Herrera Dávila y Alvear, publicadas en Sevilla quince años antes de aquel primer referente, aunque es menos extenso y no haga referencias explícitas a otros autores y obras, contiene un conjunto de elementos que son comunes a la obra de Ibáñez como veremos después.

Las dos características comunes más destacadas son sin duda la concepción amplia que ambas obras tienen de la Estadística, cuyo detalle analizaremos más adelante, y el papel tan relevante que ambas obras confieren a la Estadística como herramienta útil para los gobiernos de los Estados y para los propios individuos. Esta última motivación práctica impulsó el desarrollo de la ciencia estadística, de ahí que hayamos considerado conveniente realizar un breve análisis de las estadísticas más relevantes que se confeccionaron en España precedentes a la primera mitad del siglo XIX.

En la segunda parte de este trabajo analizaremos más extensamente el contenido del pequeño manual de Herrera Dávila y Alvear.

Precedentes

Según Sanz Serrano con el reinado de Felipe V se vuelven a impulsar los trabajos estadísticos y geográficos así como las *informaciones* tales como las descripciones de los Corregimientos de Tarragona, Tortosa y Mataró (entre 1712 y 1728). En 1770 el Consejo de Castilla emite una circular a los pueblos de España una de cuyas consecuencias es la formulación por Francisco Mariano Ninfo de un cuestionario que contiene 13 preguntas sobre agricultura, 13 sobre manufacturas, fábricas y artes, 19 sobre comercio, 7 sobre las ciencias, 4 sobre política y gobierno, 4 sobre noticias eclesiásticas, 7 sobre salud pública y 8 sobre Historia Natural. Por otra parte, Tomás López, geógrafo real, reunió una multitud de datos que están recogidos en 20 volúmenes en folio en la Biblioteca Nacional. En 1749 se inicia la realización del Catastro del Marqués de la Ensenada que no tenía, según Sanz Serrano, más antecedente en España que las *Relaciones Topográficas y Estadísticas* ordenadas por Felipe II. Los trabajos debieron prolongarse durante siete u ocho años. El cuestionario constaba de 40 preguntas, que fueron recibidos por las autoridades de los pueblos, comprende cuestiones demográficas, geográficas, económicas y judiciales sobre las tierras, cultivos, derechos impuestos sobre las tierras, ganado, número de vecinos, establecimientos mercantiles, ocupaciones, etc.

Durante el reinado de Carlos III se siguen produciendo trabajos estadísticos de gran importancia como el *Censo del Conde de Aranda* de 1768 arrojando una población total para España de 9.307.804 personas. En 1787 Floridablanca realiza un nuevo censo que da un total de 10.609.879 habitantes.

En el reinado de Carlos IV se emprende la realización del *Censo de Frutos y Manufacturas de España e Islas adyacentes, con reflexiones importantes sobre la estadística de cada una de las provincias hecho por el Departamento de Fomento General del Reino y de la Balanza de Comercio* (1799).

La primera mitad del siglo XIX en España, como es sabido, está políticamente repleta de vicisitudes y vaivenes. En el ámbito de las estadísticas existe una gran cantidad de normativa que regula las actividades estadísticas oficiales. Así, por ejemplo, en 1813 se encomienda a los Ayuntamientos la creación de un Registro de nacimientos, matrimonios y defunciones cuya información remitían periódicamente a instancias administrativas supramunicipales. Sanz Serrano dice de esta época:

> *"Bien es cierto, que las noticias que empezaron a solicitarse desde esta época, sin descanso ni tregua, no todas tenían carácter estadístico en el sentido técnico de la palabra; pero al reclamarlas se reveló el propósito de conocer a España y se demostró con toda evidencia que, para gobernar con acierto, era necesario conocer a fondo todos los objetos de la administración estatal"*[1]

Una aportación interesante del primer tercio del siglo XIX es la de Álvaro Flórez Estrada, Intendente Militar de Andalucía y autor de una obra titulada *Plan para formar la Estadística de Sevilla* (1814). Este autor tiene conciencia de la gran importancia que tiene la Estadística en al administración y gobierno de los pueblos así como de su prosperidad. Por otra parte, concibe la Estadística desde una perspectiva amplia como se muestra en la división que hace de su libro. Distingue los siguientes capítulos:

- Estadística de la agricultura y de la tierra

[1] SANZ SERRANO, A.: *Resumen Histórico de la Estadística en España.* María Gómez Ediciones. Madrid, 1956, p. 119.

- Estadística de ganados
- Estadística de las casas y demás edificios
- Estadística de la Población
- Estadística de la Industria y del Comercio
- Estadística de las Festividades
- Estadística de la Educación

Siendo ministro de Hacienda Mateo Miguel de Ayllón en 1843, se creó una Comisión de Estadística bajo la presidencia de Pascual Madoz, cuyas funciones principales eran reunir información estadística de los Archivos españoles desde el siglo XIV y elaborar un proyecto de Ley para la creación de un cuerpo administrativo con atribuciones estadísticas. Por otra parte, Madoz impulsa la formación de dos jóvenes (Juan Bautista Trúpita y José Magán) en el estudio de la estadística, para lo que dichos jóvenes son enviados a Francia, Bélgica e Inglaterra.

En 1844 se crea la primera Cátedra de Estadística por la Sociedad Económica Matritense, que es ocupada como ya se ha dicho por José María Ibáñez, entonces secretario de la Comisión de Estadística.

En ese mismo año Fermín Caballero publica el *Manual Geográfico-Administrativo de la Monarquía Española*, obra en la que se recogen descripciones geográficas y administrativas de la España de la época.

En 1845 comienza la impresión del *Diccionario Geográfico, Estadístico e Histórico* en 16 tomos en folio y finaliza en 1850.

En 1856 por un Decreto de 3 de noviembre se encarga a la Comisión de Estadística la tarea de formar la Estadística General del Reino y que representa el primer precedente importante de lo que con el paso del tiempo constituye el INE y la Estadística Oficial.

Las *lecciones de estadística* de Herrera Dávila y Alvear

Hasta ahora no nos ha sido posible encontrar alguna información reseñable de carácter biográfico de los autores de este trabajo. Sabemos que sus nombres aparecen en distintas obras de carácter divulgativo, desde sus *Lecciones de Antigüedades Romanas (1827), Lecciones de Heráldica (1830), Lecciones de Jeografía de España (1829), Lecciones de Historia de los Imperios Antiguos (1829), Lecciones de Economía Política (1827), etc.*, hasta las *Lecciones de Estadística (1829)* que nos ocupan especialmente[2]. Tenemos constancia de que al menos algunos de estos títulos fueron encargados a otros autores para formar parte de la colección de obras que dirigían Herrera Dávila y Alvear.

Existe un ánimo explícito de realizar una gran colección enciclopédica de obras de todos los campos del conocimiento como se puede apreciar en la planificación que se hace de las mismas al final de las *Lecciones* que tratamos aquí. Esta obra constituye, según se desprende de una nota final, del cuaderno 13. Las *Lecciones de Jeografía* ocupa el número 16.

Por otra parte, tenemos noticia de la existencia de una obra titulada *Guía de Sevilla* cuyo autor es Herrera Dávila, que aún no hemos podido localizar, en la que el autor afirma que

[2] La obra completa pretendía cubrir una amplia temática, estaba previsto la publicación de 127 títulos de los cuales sólo tenemos constancia de la publicación de 16 de ellos. Incluso se anunciaba un tomo dedicado al Cálculo de Probabilidades. No sabemos por qué circunstancia se interrumpió la edición de estos tomos.

existen 96683 habitantes en 1832 en la ciudad de Sevilla. Esta información nos hace pensar que la elaboración de este trabajo fuese obra de los autores que en ella aparecen, o de al menos uno de ellos.

En 1895 Bertillon escribe en su *Curso Elemental de Estadística Administrativa* lo siguiente:

"El desarrollo de la estadística en España ha sido muy lento. Estadísticas parciales se han hecho aquí en todas las épocas, lo mismo que en todas partes; pero informaciones generales y uniformes sólo existen desde hace poco. El primer censo data de 1860"[3].

Omitimos cualquier comentario sobre el contenido de la cita, pero lo que sí destacamos es que esta es la imagen que desde fuera se tiene y que explica las quejas y lamentos de algunos autores españoles de la época. Actualmente no se discute la siguiente afirmación que tomamos de Sánchez-Lafuente:

"No podemos eludir comenzar hablando de estadísticas y no de Estadística, aunque nuestro trabajo lleva por título el análisis histórico de la ciencia, pues, aparte de razones que son obvias, aunque no dejaremos de enunciarlas, España, entendemos, se anticipó al resto de Europa en esta clase de trabajos"[4].

Como ya se ha dicho, los gobiernos de la época otorgan un papel crucial a la Estadística en la persecución del objetivo de conseguir la felicidad de los pueblos. Flores Estrada escribe en 1814:

"El objeto de la Estadística es el de que el Gobierno y sus ciudadanos tengan conocimiento fácil y pronto de todos los datos que sean necesarios para formar con acierto y seguridad los planes que puedan conducir a mejorar la suerte de los pueblos"[5].

Ibáñez en la introducción de su tratado destaca la importancia de la ciencia estadística pues *"es considerada ya en todas las naciones cultas como la más esencial para la resolución de muchas cuestiones que interesan altamente al bienestar de los pueblos"*[6].

Entre ambas referencias merece la pena destacar la relación que tiene la Estadística con la felicidad de la nación según los autores de las *Lecciones de Estadística*, referencia por otra parte común en los escritos que aparecen en las obras de la época como queda mostrado en las citas previas. Después de realizar una enumeración de actividades que conducen a la felicidad de una nación respecto a cuatro elementos principales como sus habitantes, sus tierras, sus productos y su moneda circulante en dichas *Lecciones* se dice que *"tales son los principales medios de hacer feliz á una nacion, y estos son tambien los objetos de que se ocupa la Estadística"*.[7]

La concepción que tiene Ibáñez de la Estadística es amplia (comprendería lo que hoy podríamos llamar grandes sistemas de información geográfica, económica, moral, jurídica, etc.) y se destaca la gran utilidad que tiene no sólo para los gobiernos de los Estados sino para

---

[3] BERTILLON, J.: *Curso Elemental de Estadística Administrativa.* Hijos de Reus Editores. Madrid, 1907, p. 39.

[4] SÁNCHEZ-LAFUENTE FERNÁNDEZ, J.: "Historia de la Estadística como ciencia en España (1500-1900)". *Estadística Española, núms. 58 y 59. Madrid, 1973, p. 20.*

[5] Ver SÁNCHEZ-LAFUENTE FERNÁNDEZ, J.: Ibidem, p. 113.

[6] IBÁÑEZ, J.M.: *Tratado Elemental de Estadística.* Tomo I. Reimpresión INE, Madrid, 2006, p. IV.

[7] HERRERA DÁVILA, J. Y ALVEAR, A.: *Lecciones de Estadística.* Imprenta de Mariano Caro. Sevilla, 1829, p. 7.

otras instituciones e incluso para los mismos individuos. Siguiendo a Gioja (*Filosofia Della Statistica,* Milano, 1826) Ibáñez nos dice:

> *"La Estadística pues, al examinar y describir su situación topográfica, sus límites, sus disposiciones naturales, la forma de su suelo, el clima etc., facilita al gobierno los conocimientos más necesarios para la mejor construcción de los caminos y canales, para la navegación de los ríos que sean susceptibles de este uso, para conservar los montes y bosques, beneficiar minas, abrir puertos militares o de comercio, designar puntos de defensa, situar fortalezas, fijar la estación de las tropas de mar y tierra, arreglar sus movimientos etc., etc.; en una palabra para determinar con acierto sobre todos aquellos objetos en cuya buena o mal dirección y economía ha de influir forzosamente la situación, clima y naturaleza del suelo"*[8].

En cuanto al objeto de la Estadística nos dicen Herrera Dávila y Alvear ya en el mismo prólogo de sus *Lecciones de Estadística*:

> *"Ella reúne los datos sobre que han de girar en gran parte los cálculos de la política: con las investigaciones del número de habitantes, fertilidad del suelo, fábricas, rentas, fuerza militar y otras circunstancias presenta una multitud de conocimientos, que no sólo son útiles á los encargados de la administracion pública para hacer felices á los pueblos, sino tambien á los simples particulares para sacar en su favor inmensas ventajas de los progresos de la agricultura, comercio, artes y toda especie de industria, al paso que despleganen la sociedad estos elementos de prosperidad pública."*[9].

Es también interesante destacar que utilizan una definición del tipo de la escuela alemana, al decir en el prólogo que la Estadística es el *Arte de formar el inventario de un Estado*.

Salvo los libros de Peuchet (*Essai d´une statistique générale de la France*, 1798) y de Gioja (*Filosofia della Statistica,* 1825), no se habían publicado los libros que fueron referencia en el tratado de Ibáñez. Tal vez las fuentes dignas de mención que sospechamos hayan podido utilizar los autores, aunque no citan a ningún autor ni a ninguna obra como ya se ha dicho, sean éstas y la obra de Achenwall, bien directamente o a través de sus referencias en la obra de Gioja.

Existe también un interés explícito en estos primeros momentos en que aún no está delimitado el objeto de la nueva ciencia de la Estadística por distinguir la Estadística y la Aritmética Política que según Ibáñez *"difieren enteramente de esta por los medios de que se vale, por su manera de proceder, y por los principios mismos en que se funda sus investigaciones y sus cálculos"*[10]. Siguiendo a Peuchet, nos dice Ibáñez que la *"Estadística no puede admitir, ni de ellas hace uso para los suyos por no poderlas considerar exactas, pues que la exactitud en los datos y en las bases para los cálculos, rechazando toda suposición arbitraria, es un principio constante en la Estadística"*[11].

También en el breve prólogo de las *Lecciones* se destaca esta idea que, como hemos dicho, representa una constante en muchas de las obras dedicadas a la Estadística. Dan por sentado que la Aritmética Política es algo distinto, aunque muestra el papel relevante que esta ciencia tiene en su aplicación sobre los datos estadísticos:

---

[8] IBÁÑEZ, J.M.: Op. Cit., pp.43-44.
[9] HERRERA DÁVILA, J. Y ALVEAR, A.: *Op. Cit.*, pp. 3-4.
[10] IBÁÑEZ, J.M.: Op. Cit., p.53.
[11] IBÁÑEZ, J.M.: Op. Cit., p.53.

> *"Completamos su doctrina con dos lecciones de Aritmética política, tanto para dar idea de este cálculo naciente entre nosotros, como para demostrar, por medio de alguna aplicacion de los datos estadísticos, la importancia de la ciencia que los suministra"*[12].

Existe, por otra parte, un claro interés pedagógico al considerar que puede prestar un *"servicio transcendental a la juventud estudiosa y sus directores"*. La misma estructura del libro en la que un conjunto de lecciones van seguidas de unas preguntas sobre las mismas, tiene un claro objetivo pedagógico al proporcionar comodidad a padres y maestros.

El inventario de las cosas del Estado, para estos autores, se centra en cuatro elementos principales:

- Habitantes
- Tierras
- Productos
- Moneda circulante

Otro rasgo distintivo de las *Lecciones* va a ser la continua preocupación de sus autores por sistematizar la ciencia Estadística, en la que aparecerán múltiples clasificaciones y un interés por denominar con nuevos nombres de origen latino cada una de las partes en que se estructura la Estadística. Distinguen tres tipos de operaciones estadísticas:

Descriptivas: se refieren al territorio, los habitantes y las producciones de la naturaleza en general.

Positivas: se ocupa del cultivo intelectual, económico, moral y político de la población.

Aplicada: reúne los datos de las anteriores y las adapta a las necesidades y relaciones legislativas, administrativas, económicas y diplomáticas del Estado.

A la primera la llama Corografía, a la segunda Etnografía y a la tercera Nomografía. A su vez cada una de ellas comprende distintos objetos, lo que da lugar a una concepción extensa de la Estadística. Así nos dicen estos autores:

> *"24. La Corografía comprenderá la situación geográfica, el suelo y el clima, las producciones de la naturaleza y las habitaciones de los hombres.*
> *25. La Etnografía expondrá el estado de la población, de la agricultura, de la industria y de la civilización.*
> *26. La Nomografía discurrirá sobre las leyes, las administraciones, la economística y la diplomacia."*[13].

No obstante, cuando desarrollan la Etnografía nos dicen que sus elementos constituyen *"el verdadero núcleo de la ciencia estadística"*[14]

Atendiendo a la mayor o menor extensión que se quiera dar a las investigaciones, establecen también la siguiente división de la Estadística:

- Analítica

[12] HERRERA DÁVILA, J. Y ALVEAR, A.: Op.cit., p. 4.
[13] HERRERA DÁVILA, J. Y ALVEAR, A.: Op.cit., p. 11.
[14] HERRERA DÁVILA, J. Y ALVEAR, A.: Op.cit., p. 38.

- Especial
- Interna

La primera toma como objeto todos los Estados, la segunda se limita a un solo Estado y la tercera se refiere a una parte de un Estado. Se trata de una clasificación que sigue un criterio de extensión del objeto sometido a estudio.

Siguiendo con este interés pedagógico de esta obra destacan los autores seis reglas principales que debe cumplir cualquier operación estadística, que nosotros hemos dado nombre a partir de la exposición del contenido de cada una de ellas. En estas reglas se vuelve a apreciar el interés práctico que tienen los autores en las operaciones estadísticas, elemento por otra parte clave para obtener informaciones útiles y veraces que ayuden en el gobierno de los Estados.

1. Coherencia: entre las ideas del estadístico y las características morales y económicas del Estado.
2. Rigor: se deben apoyar en "*hechos y cómputos*" verdaderos.
3. Orden: debe existir una disposición clara y ordenada de la información.
4. Sencillez: se deben evitar exposiciones prolijas y complejas.
5. Utilidad: se deben recoger aquellos datos que sean significativamente útiles para el Estado.
6. Comprensión: debe proporcionar herramientas que permitan el conocimiento de las cosas del Estado.

Las fuentes pueden ser públicas (edictos, estatutos, registros públicos, catastros, censos de población, etc.) o privadas (mapas, gacetas, periódicos, obras impresas o manuscritos, etc.).

A partir de la lección IV recogen de un modo sistemático el contenido de cada una de las partes de las operaciones estadísticas, recordemos: corografía, etnografía y nomografía.

Para cada una de estas partes establece varias subdivisiones y en cada una de estas a su vez define un conjunto de tablas. En total diseñan el contenido de 130 tablas. Se vuelve a observar un gran sentido práctico, es decir sus lecciones están orientadas a la práctica de elaborar estadísticas de un modo riguroso y sistemático. Veremos después cómo, a modo de ejemplo, proponen cuestionarios relacionados con algunas de las 130 tablas que describe.

La corografía la divide en cuatro apartados: sitio geográfico, clima y suelo, producciones naturales y habitaciones de los hombres. Llama la atención la amplitud de los contenidos que en cada subdivisión se detallan como objeto de las tablas que la componen. Representa un verdadero interés por conocer exhaustivamente las cosas del Estado.

La Etnografía la divide en cuatro partes: Población, agricultura, industria y civilización.

La Nomografía la divide en cuatro partes: legislación, administración, economística y diplomática.

Llega a un detalle que resulta especialmente interesante desde un punto de vista práctico. A título orientativo establece un conjunto de ítems o preguntas que pueden servir para el estadístico. De cada una de las tablas de la subdivisión "*sitio geográfico*" que se incluye dentro de la Corografía, establece en torno a una veintena de ítems o cuestiones que guíen al estadístico para elaborar cada una de dichas tablas.

Dedican, por último, dos lecciones a la Aritmética Política cuyo objetivo es aplicar el cálculo a los objetos de la política tales como el valor del terreno, el número de habitantes de un país, la duración de la vida humana, el cómputo de los rendimientos de las riquezas presumidas de un Estado, las proporciones entre las diversas clases de individuos que componen la sociedad y cualesquiera estimación fundada sobre cálculos fijos o aproximados. Luego sólo se ofrecen ejemplos exclusivamente de demografía.

El carácter pedagógico del libro provoca que dediquen los autores un tema final que contiene 136 preguntas sobre las lecciones.

Conclusiones

Con independencia de cuál sea la verdadera autoría del pequeño manual que acabamos de reseñar está claro que es el primero que se edita por españoles adelantándose en 15 años al conocido Tratado de José María Ibáñez. Como todos los primeros tratados, mostraba un interés desmesurado por integrar en esta nueva ciencia el mayor número de contenidos sobrepasando los objetivos que hoy día se plantean, pero era para demostrar y convencer a políticos y público en general de su importancia, por lo que se invadían los contenidos típicos del Derecho, la Geografía, etc. A pesar de su bisoñez este manual sí plantea la idea de la integración de las diferentes investigaciones estadísticas y la necesidad de coordinar los distintos trabajos ordenándolos al fin común del conocimiento de las cosas del Estado.

Bibliografía

ACKERKNECHT, ERWIN.: "Hygiene in France, 1815-1848", *Bulletin of the History of Medicine*, vol. XXII, 1948.

BOULLET, M.N. : Dictionnaire Universel des Sciences, des Lettres et des Arts, Paris, Hachette, 1874.

BOURGET, MARIE NOELLE. : *Déchiffrer la France. La Statistique Départamentale à l'Epoque Napolénoienne*, Paris, Editions des Archives Contemporains, 1989.

COLEMAN, WILLIAM.: Death is a Social Disease. Public Health and Political Economy in Early Industrial France, The University of Wisconsin Press, 1982.

CHÁZARO, LAURA.: Midiendo el cuerpo de una nación. Ensayo sobre la estadística médica en México a finales del siglo XIX, tesis doctoral en Filosofía, Facultad de Filosofía y Letras, UNAM, México, 2000.

DESROSIÈRES, ALAIN. : La Politique des Grands Nombres. Histoire du Raison Statistique, Paris, Editions La Découverte, 1993.

DESROSIÈRES, ALAIN. : "¿Cómo fabricar cosas que se sostienen entre sí? Las ciencias sociales, la estadística y el Estado", *Archipiélago. Cuaderno de crítica de la cultura. Primera Carpeta*, Madrid, núm. 20, 1995.

DUPAQUIER, JACQUES.: *L'Invention de la Table de Mortalité*, Paris, PUF, 1996

ESCALANTE GONZALBO, FERNANDO.: *Ciudadanos imaginarios*, México, El Colegio de México, 1992.

GORBACH, FRIDA.: "El placer del encierro: La imagen de un hermafrodita", *Argumentos. Estudios críticos de la sociedad*, núm. 29, UAM-X, abril de 1998.

GUERRA, FRANCOIS, XAVIER. Y A. LEMPÉRIERE ET AL.: Los espacios públicos en Iberoamérica. Ambigüedades y problemas, siglos XVII-XIX, México, FCE, 1988.

GUERRA, FRANCOIS XAVIER.: *MÉXICO: Del antiguo régimen a la revolución*, México, FCE, t. 1, 1988.

HACKING, IAN.: *La domesticación del azar*, Madrid, Gedisa, 1991.

HARRIS, RUTH.: Murders and Madness: Medicine, Law and Society in the Fin de Siècle, Oxford, Clarendon Press, 1991.

HIDALGO Y CARPIO, LUIS Y GUSTAVO RUIZ Y SANDOVAL.: *Compendio de medicina legal arreglado a la legislación del Distrito Federal*, t. I, México, Imp. de Ignacio Escalante, 1877.

HOBSBAWM, ERIC.: *Naciones y nacionalismo desde 1780*, Barcelona, Crítica, 1991.

LA BERGE, ANN F., Mission and Method. The Early Nineteenth-Century-French Public Health Movement, Cambridge, Cambridge University Press, 1992.

LA BERGE, ANN F.: "The Early Nineteenth Century French Public Health Movement: The Disciplinary Development and Institutionalization of Hygiene Publique", *Bulletin of the History of Medicine*, vol. 58, numb. 3, 1984.

LA VERGATA, ANTONELLO.: "Biology and Sociology of fertility. Reactions to the Malthusian threat. 1798-1933", en Brian Dolan (edit.), *Malthus, Medicine and Morality. 'Malthusianism' after 1798*. Clio Medica 59, Ámsterdam, Rodopi, 2000, pp. 189-210.

LAVISTA, RAFAEL.: "Relaciones entre la medicina y la jurisprudencia", *Concurso Científico*, México, Oficina Tipográfica de Fomento, 1895.

LICEAGA, EDUARDO.: "Algunas consideraciones acerca de la higiene social en México", *Concurso científico y artístico del Centenario, Academia Mexicana de Jurisprudencia y Legislación*, México. Tip. Vda. de F. Díaz de León, Sucs., 1911.

LIER, E.: "Ginecología. La esterilidad en los matrimonios", *Gaceta Médica de México*, tomo XXV, núm. 12, 1890, pp. 221-240..

LOBATO, JOSÉ GUADALUPE.: "Higiene. Sociología en sus relaciones con la demografía y demología mexicanas", *Gaceta Médica de México*, tomo XV, núm. 16, 1880, pp. 357-371.

LÓPEZ BELTRÁN, CARLO.: "'Les maladies héréditaires' 18th Century Disputes in France", *Revue of the History of Science*, XVII, 3, 1995.

MALANCO, FERNANDO.: "Fisiología psicológica. Conexión entre lo físico y lo moral del hombre. Ventajas que de ella puede sacar la medicina", *Gaceta Médica de México*, tomo XXXIV, núm. 15, 1 de agosto de 1897, pp. 406-411.

MEJÍA, DEMETRIO.: "Estadística de la mortalidad en México", *Gaceta Médica de México*, tomo XIV, núm. 14, 15 de junio de 1879, pp. 273-301.

PICK, DANIEL.: *Faces of Degeneration. A European Disorder, c. 1848- c.*1918, Cambridge, Cambridge University Press, 1993.

PORTER, THEODORE.: "Making Things Quantitative", *Science in Context*, Cambridge, vol. 7, number 3, Autumn 1994, pp. 389-407.

PORTER, THEODORE.: *The Rise of Statistical Thinking. 1820-1900*, Princeton University Press, 1986.

QUETELET, ADOLPHE. : "Mémoire sur les Lois des Naissances et de la Mortalité à Bruxelles", S.l., *Extrait des Mémoires de l'Académie Royale des Sciences et Belles Lettres de Bruxelles*, t. III, s/d.

QUETELET, ADOLPHE. : Sur l'homme et les développements de ses facultés. Essai de Physique Sociale, Paris, Fayard, 1991.

QUINTAS ARROYO, Juan.: "Estadística médica", GMM, 1874.

RAMÍREZ, ROMÁN.: "Nociones preliminares de patología", *Resumen de la Medicina legal y ciencias conexas,* México, Oficina Tipográfica de Fomento, 1901.

REYES, JOSÉ MARÍA.: "Estadística de mortalidad en la capital, con arreglo al censo de su población. Su estado patológico. Primera y segunda parte", *Boletín de la SMGyE*, México, Imprenta del Gobierno, 1869.

REYES, J.M.: "Higiene pública. Mortalidad de la niñez", *Gaceta Médica de México*, tomo XIII, núm. 20, 11 de julio de 1878, pp. 377-385.

REYES, J. M.: "Memoria leída por el secretario del Consejo Central de Salubridad el día 17 de enero de 1867, relativa a los trabajos de esta corporación en el año próximo pasado", Documento del Fondo Reservado, Lafragua, UNAM**,** 1867.

RUIZ Y SANDOVAL, G.: Estadística de la mortalidad y sus relaciones con la Higiene y la patología de la Capital, México, 1872.

SIERRA, JUSTO.: Obras Completas. Evolución política del pueblo mexicano, México, UNAM, 1986.

STIGLER, STEPHEN.: *The History of Statistics. The Measurement of Uncertainty before 1900*, Belknap Press of Harvard University Press, 1986.

URÍAS HORCASITAS, BEATRIZ.: Indígena y criminal. Interpretaciones del derecho y la antropología en México. 1871-1921, México, Universidad Iberoamericana, 2000.

URÍAS HORCASITAS, BEATRIZ.: Historia de una negación: la idea de igualdad en el pensamiento político mexicano del siglo XIX, México, UNAM, 1996.

VILLERMÉ, LOUIS-RENÉ. : "Hygiène Publique. Mémoire sur la Mortalité dans les prisons", *Annales d'Hygiène Publique et de Médicine Légale*, t. I, Paris, 1829.

WILLIAMS, ELIZABETH.: The Physical and the Moral. Anthropology, Physiology, and Philosophical Medicine in France, 1750-1850, Cambridge, Cambridge University Press, 1994

Capítulo 3

José María Ibáñez Ramos: Primer catedrático de estadística

FRANCISCO JAVIER MARTÍN-PLIEGO LÓPEZ
Universidad Rey Juan Carlos

Introducción

D. José María Ibáñez y Ramos[1] nació en Granada el 20 de octubre de 1793, hijo de Antonio Ibáñez Navarro, turolense, y de María Francisca Ramos de la Tita, natural de Ronda, casado el día 29 de marzo de 1826 en Granada con Dña. María de la Presentación Moreno. Tuvo una hija, Dña. Emilia, y falleció en su residencia madrileña de la calle de Fomento nº 21, el 7 de octubre de 1856, víctima de un ataque de epilepsia según consta en el parte del correspondiente facultativo. Se le enterró en el cementerio de la Archicofradía Sacramental de San Ildefonso y de San Marcos[2] de Madrid.

Estudió latín, griego y francés y matemáticas puras y mixtas, física, química e historia natural en el afamado colegio granadino de D. Francisco Dalmau, obteniendo la calificación de *excelente* en los consiguientes exámenes públicos. Debió ser un aventajado alumno ya que empezó a dar clases de matemáticas en principio en el mismo colegio donde estudió y posteriormente en la Real Academia Militar de Granada. Concluidos estos estudios, en 1819, pasó a ingresar en la Universidad de la misma ciudad donde cursó estudios de filosofía y leyes, logrando, al superar las correspondientes pruebas *nemine discrepante,* el grado de bachiller en Derecho Civil, obteniendo el título de abogado de la Real Chancillería de Granada en 1826[3] y de los Reales Consejos en 1833.

[1] Véanse Archivo Histórico Nacional, fondos contemporáneos, Ministerio de Justicia, Magistrados y Jueces, legajo 4496-1, expediente 3909 y Ministerio de Hacienda, serie general, legajo 5093, expediente 73.
[2] Terreno donde en la actualidad se encuentra el Estadio Municipal de Vallehermoso.
[3] Véase certificado de su nombramiento en Archivo Histórico Nacional, fondos contemporáneos, Consejos, legajo 12102, expediente 152.

Fue Fiscal General de Correos en la Dirección General de Correos de España e Indias por nombramiento del 13 de agosto de 1834, Ministro Honorario del Consejo de Hacienda, Subdelegado y Juez Conservador de la Imprenta Nacional y Secretario de la Comisión de Estadística creada en 1843 , socio de mérito de la Sociedad de Amigos del País de Granada, su tierra natal, socio de número de la Real Sociedad Económica Matritense de Amigos del País y redactor principal de *El Amigo del País*, periódico editado por la Sociedad Económica Matritense. En 1844 fue nombrado profesor encargado de la Cátedra de Estadística establecida en esta última institución[4].

Pascual Madoz, amigo y mentor suyo, tenía una loable opinión de D. José María Ibáñez, así en este sentido dice[5]:

> "*Es el señor Ibáñez uno de esos hombres que por un exceso de modestia quieren vivir en la oscuridad, pero cuyos talentos debe aprovechar el Gobierno, cualquiera que sean las opiniones de los ministros: recomiendo a mis lectores el tratado elemental de estadística que el señor Ibáñez publica, del cual ya han salido dos interesantes cuadernos.*"

En los países europeos de nuestro entorno, en el siglo XIX, se empezaron a establecer gérmenes de lo que devendrían a ser más tarde "*oficinas centrales de estadística*", con objeto de coordinar esfuerzos y de sentar unas bases comunes a los diferentes trabajos estadísticos que hubieran lugar para facilitar su integración. En España no dejó de reflejarse esta corriente y hubo diferentes instrucciones gubernamentales a principios de este siglo para ir dirigiendo y tutelando los diferentes trabajos estadísticos que se realizaban. En 1843, el Ministro de Hacienda D. Mateo Miguel Ayllón[6] nombró una *Comisión de Estadística* bajo la presidencia de D. Pascual Madoz e Ibáñez, actuando como Secretario de dicha *Comisión* D. José María Ibáñez y Ramos. Esta *Comisión* propuso, entre otras medidas, la creación de dos Cátedras de Estadística, pero este propósito fue abandonado por falta de medios económicos. Felizmente esta idea fue recogida pocos años más tarde por la Sociedad Económica Matritense que creó una Cátedra de Estadística en Madrid de la que se encargó D. José María Ibáñez, el cual empezó sus enseñanzas en 1844 culminando su primer curso 38 alumnos. La *Comisión* fue fecunda en sus trabajos y a partir de sus propuestas se emanaron una amplia diversidad de instrucciones, reales órdenes y reales decretos.

Toda esta callada labor dio lugar a que el 3 de noviembre de 1856, siendo Presidente del Gobierno D. Ramón María Narváez, se creara por Decreto la *Comisión General Estadística del Reino,* elevándola de rango, pues se situó directamente bajo la Presidencia del Consejo de Ministros. En dicho Decreto se dice literalmente lo siguiente:

---

[4] Cuando el INE me encargó el estudio preliminar para la edición facsímil del Tratado Elemental de Estadística de J.M. Ibáñez realicé investigaciones para ofrecer una breve reseña biográfica de este autor encontrando por mi cuenta algunos datos de interés. Simultáneamente el funcionario del INE Antonio Laborda hizo sus pesquisas hallando más información sobre este autor que recogí en mi estudio preliminar queriéndole citar para reconocer al sr. Laborda su trabajo pero me dijo que no hacía falta y que no tenía ningún documento suyo al cual pudiera referirme. Posteriormente a mi entrega de este estudio (Septiembre de 2006) el sr. Laborda publicó el opúsculo que reseño en la bibliografía en donde no me reconoce la autoría de las pesquisas que hice. Como investigador académico quiero dejar constancia de este hecho.

[5] Véase MADOZ, P: Diccionario Geográfico-Estadístico-Histórico de España y sus posesiones de ultramar. Tomo I. pág. XIX. Madrid. 1848.

[6] Véase SANZ SERRANO; A: Resumen Histórico de la Estadística en España. págs. 126-133. INE. Madrid. 1956

"Todos los gobiernos anteriores se han ocupado con asiduidad de trabajos estadísticos; pero olvidando unos el objeto, circunscribiéndose otros a determinadas clases, y finalmente, verificándolo los demás aisladamente, aunque con esmero, ha venido a encontrarse el actual con grandes vacíos, que no pueden llenarse con trabajos ejecutados en épocas diversas y que es preciso darles conexión y unidad indispensable para que produzcan el resultado apetecido.

Los trabajos parciales emanados de algunos Ministerios tienen un mérito indisputable, principalmente los que se refieren a la Administración rentística, y que atestiguan a cada periodo los progresos que va haciendo en nuestro país la ciencia administrativa; pero, falta a la Estadística española, para que la ciencia, el Gobierno y los pueblos obtengan los resultados que son de desear, que los trabajos estadísticos y su dirección sean uniformes; que partan de un mismo centro que les dé impulso comunicándoles el orden y la relación que deben tener entre sí, y que las bases de las investigaciones sean perfectamente determinadas y se ejecuten sin los embarazos que la errónea opinión de los pueblos o los recelos del fisco pudieran crear".

Por desgracia D. José María Ibáñez y Ramos, a pesar de haber colaborado activamente, con sus estudios, docencia y actividad en la Administración, para que se llegara a este momento histórico, no lo pudo ver pues murió el 7 de octubre de ese mismo año 1856. De no ser por su fallecimiento seguramente también habría estado en esta nueva *Comisión.*

De los trabajos de José María Ibáñez relativos a la incipiente ciencia de la Estadística destacaremos los siguientes.

El Tratado Elemental de Estadística

Con objeto de que sus alumnos pudieran seguir mejor sus explicaciones publicó el manual *Tratado Elemental de Estadística, así en la parte filosófica y de teoría, como en la aplicación de sus principios a la práctica*[7]. La publicación de este tratado[8] se realizó en fascículos según se desprende del anuncio inserto en la ya citada revista *Amigos del País* en el que se indica que el Tratado " *constará de dos tomos regulares, inclusas las tablas: publicándose por entregas de 54 páginas en octavo prolongado, a 5 reales en Madrid, y a 6 en las provincias, franco el porte y anticipándose el importe de cada entrega. La primera saldrá a mediados de enero actual y se darán sucesivamente dos en cada mes, si es posible. Al publicarse la tercera quedará cerrada la suscripción y se venderán a 6 reales*".

Dentro del escaso conocimiento que se ha tenido sobre este personaje, hay que reseñar que en lo más conocido de él, que es esta obra, también el conocimiento es incompleto pues la mayoría de los historiadores de la estadística cuando se refieren al Tratado citan exclusivamente su primer Tomo, como si no se hubiera publicado nunca el segundo, e incluso otros niegan abiertamente su existencia[9].

---

[7] Editado en dos tomos, el primero publicado en 1844 y el segundo en 1845 en la Imprenta del Colegio de Sordo-mudos de la calle del Turco nº 11 de Madrid, establecimiento colindante con la sede de la Sociedad Económica Matritense. La calle del Turco es la actual vía madrileña del Marqués de Cubas.

[8] En noviembre de 2006 y con motivo del 150 aniversario de la Estadística Oficial en España, el I.N.E. realizó una edición facsimilar de los dos tomos de este texto.

[9] Véase SÁNCHEZ-LAFUENTE FERNÁNDEZ, J. : Historia de la Estadística como ciencia en España (1500-1900). INE. Madrid. 1975. y VELASCO PÉREZ, R. : " Una nota sobre José María Ibáñez, primer catedrático

No obstante, este texto tiene el indudable mérito, entre otros, de ser el primer manual completo y detallado sobre estadística publicado en castellano puesto que aunque anteriormente se habían publicado dos opúsculos, uno de D. Francisco Dalmau[10] y otro de los publicistas José Herrera Dávila y Antonio Alvear[11], éstos ni tenían la extensión en contenido, ni abarcaban toda la temática de los dos tomos del manual de Ibáñez. José María Ibáñez conoce y cita[12] la publicación de Dalmau, profesor suyo en sus estudios de matemáticas como antes reseñamos, que fue el resultado de una investigación estadística en la provincia de Granada donde muy probablemente participara en sus trabajos de campo y lo refleja, al hablar del método de la medición de los terrenos, al comentar que el método que propone "*no sería tampoco nuevo en España*" ya que en 1819 el profesor Dalmau fue encargado por el Gobierno de formar la Estadística general de Granada y usó el método del levantamiento previo del plano topográfico del territorio granadino por medio de técnicas matemáticas exactas.

Los otros manuales que fueron utilizados como fuente de inspiración por los estadísticos universitarios de la segunda mitad del siglo XIX fueron traducidos y editados en fechas posteriores al de José María Ibáñez. El primero fue el de Pierre Armand Dufau que se editó[13] en Madrid en 1845 y el segundo texto clásico de estadística traducido[14] fue el de Alexandre Moreau de Jonnes editado en 1857. Así, pues, el texto de Ibáñez junto a estos dos recién citados se convirtieron en las fuentes principales que se emplearon para la confección de los manuales que se usaron como texto en las cátedras de Economía Política y Estadística o de Geografía y Estadística de las Facultades de Derecho en la última mitad del diecinueve[15].

José María Ibáñez también se inspiró, fundamentalmente, en tres textos, sobre todo en el de Dufau, claro está en su versión francesa original[16], y además cita y se apoya repetidamente en el de Jacques Peuchet[17] y en el de Melchiorre Gioja[18]. De los dos primeros recoge la idea de la *aritmética política* de la escuela inglesa es algo diferente a la

---

de Estadística en España" en Homenaje al Profesor Juan Sánchez-Lafuente. pág. 344. Facultad de Ciencias Económicas y Empresariales de la Universidad de Málaga. Málaga. 1990.

[10] Véase DALMAU, F. : Ensayos de Estadística practicados en la provincia de Granada. Imprenta de Ibarra. Madrid. 1820.

[11] Véase HERRERA DÁVILA, J. ; ALVEAR, A. : Lecciones de Estadística. Imprenta de D. Mariano Caro. Sevilla. 1829.

[12] Véase IBÁÑEZ, J.M. : op. cit. Tomo II, pág. 332.

[13] Véase DUFAU, P.A.: Tratado de Estadística o Teoría del estudio de las leyes según las cuales se desarrollan los hechos sociales seguido de un ensayo de estadística física y moral de la población francesa. Imprente y Librería de Don Ignacio Boix. Madrid. 1845.

[14] Véase MOREAU DE JONNES, A. : Elementos de Estadística. Principios generales de esta ciencia, su clasificación, metodo, operaciones, diversos grados de certidumbre, errores y progresos, con su aplicación a la comprobación de los hechos naturales, sociales y políticos, históricos y contemporáneos. Imprenta de Francisco Abienzo. Madrid. 1857.

[15] Puede comprobarse las citas que del manual de José María Ibáñez se hace en: CARRERAS Y GONZALEZ, M; PIERNAS Y HURTADO, J.M. : Tratado Elemental de Estadística. pág. 6. Imprenta y Librería de Miguel Guijarro. Madrid, 1873. ; SALVÁ, M.: Tratado Elemental de Estadística. págs. 23 y 74. Agustín Jubera. Madrid. 1881. ; POU Y ORDINAS, A. J. : Curso de Estadística. pág. 70. Imprenta de la Viuda e H. de J. Subirana. Barcelona. 1889. y PIERNAS HURTADO. J. : Tratado Elemental de Estadística. págs. 14-15. Librería de Victoriano Suárez. Madrid. 1897.

[16] Véase DUFAU, P.A. : Traité de statistique, ou Théorie de l´étude des lois d´après lesquelles se développent les faits sociaux, suivi d´un essai de statistique physique et morale de la population française. Delloye. Paris. 1840.

[17] Véase PEUCHET, J. : Essai d´une Statistique Génerale de la France. Chez Testu, imprimeur-libraire. Paris. 1798.

[18] Véase GIOJA, M. : Filosofia della Statistica. Presso Giovanni Pirotta. Milano. 1826.

estadística, relegando, pues, a la estadística a la primera etapa de su desarrollo, es decir no anticipa que no sólo la estadística debe proceder al recuento y enumeración de las cosas objeto de estudio sino que partiendo de algún tipo de información es válida cualquier técnica fundamentada de estimación. Así, dice Ibáñez[19], parafraseando a Peuchet:

"La Estadística difiere tambien de la aritmética política. Esta no procede en sus operaciones por via de análisis: no trata de obtener los resultados por la enumeración de los objetos: sustituye el cálculo á estos medios, y de un dato más o menos probable ó cierto saca una consecuencia que establece como hecho.

Asi es que, por el conocimiento de la cantidad de granos consumidos anualmente en un pais, la aritmética política llega á saber el número de sus habitantes; porque es evidente que si algunos datos particulares hacen conocer el consumo anual de un individuo, se tendrá el número total de consumidores dividiendo la cantidad consumida anualmente por la que consume cada individuo en particular."

Unos pocos años más le hubieran hecho cambiar de opinión a D. José María, puesto que más tarde nadie puso en duda que los aritmeticopolíticos ingleses no fueron sino los precursores de la Estadística en su país, y además el uso de informaciones indirectas, relacionadas con el fenómeno principal observado, también fue propuesto por el propio Ibáñez al sugerir como un método de estimación de la población la utilización del número de bautismos, dato que prefería a los indicadores indirectos del número de matrimonios o de defunciones que los consideraba menos fiables.

Otra curiosidad que se repite innumerables veces a lo largo de los dos tomos de la obra de Ibáñez es el uso del vocablo *estadista* para designar a la persona preocupada u ocupada en los trabajos estadísticos, es decir el uso de *estadista* por el uso generalmente aceptado en la actualidad de *estadístico*. Hoy en día un estadista es un hombre de Estado, un político relevante, sin embargo la Real Academia de la Lengua Española admite como acepción, alternativa a la anterior, la de estadista como descriptor de la población, riqueza y civilización de un pueblo, provincia o nación. Por tanto, aunque, semánticamente, pudiera parecer incorrecto el uso de estadista por estadístico, sólo es un arcaicismo por otra parte admisible.

También hoy día llamaríamos mejor *principios metodológicos* a lo que Ibáñez, siguiendo a sus fuentes inspiradoras ya citadas, llama *principios filosóficos* de la estadística pues lo que Ibáñez pretende en su obra es fundamentalmente sentar las bases metodológicas de un completo, diríamos que casi exhaustivo, *plan estadístico* para nuestro país, según veremos cuando expongamos el contenido de esta extensa obra. Sus principios filosóficos se traducen de manera inmediata a reglas, consideraciones, recomendaciones, observaciones que han de tenerse en cuenta, es decir metodología para la obtención de datos estadísticos.

Por otra parte, después de reseñar las definiciones sobre lo que es la Estadística de Achenwall, Dufau y Peuchet, entre otros, Ibáñez establece que[20]:

"la Estadística tiene por objeto manifestar el verdadero estado de todos los elementos que constituyen la existencia física, política, moral y económica de las naciones en una época determinada.

---

[19] Véase IBÁÑEZ, J.M. : op. cit. Tomo I . págs. 74-75.
[20] Ibidem. págs. 52-53.

Definida así la Estadística se observará desde luego que no puede haber hecho ni objeto alguno dentro del Estado, que no se sujete á las investigaciones de aquella, y cuyo conocimiento interese al gobierno, á los pueblos, ó á los individuos, que es el fin esencial de las mismas investigaciones...

No parece pues, que sea preciso comprender en esta definición los medios de que ha de valerse para efectuar sus investigaciones y observaciones, ni tampoco el modo de manifestar los resultados de ellas, tanto porque los medios habrán de variarse y modificarse según las circunstancias, cuanto porque en la manera de presentar el resultado de las investigaciones y de los cálculos, pueden seguirse métodos diversos con arreglo tambien a las circunstancias de los objetos sobre que recaigan, y según su índole y naturaleza,.... "

Pero más adelante destaca que la obtención de datos de la realidad social, de los números, como dice él, como resultado de todos los hechos observados en un periodo más o menos largo sirve para resaltar las regularidades y relaciones entre los hechos sociales, civiles y económicos que en su consideración aislada y sin enlace con los demás hechos de la misma especie no se aprecian, y que una vez establecidas estas leyes se pueden deducir consecuencias muy precisas y útiles para la dirección y gobierno de un país.

El contenido de los dos tomos está dividido inicialmente en dos partes, la Parte Primera, que figura incluida íntegramente en el Tomo I y que titula *"Filosofía de la ciencia estadística, y teoría de sus principios"*. La Segunda Parte, que denomina *"Aplicación de los principios de la ciencia estadística á la práctica de sus investigaciones"*, se reparte entre los Tomos I y II. Cada una de estas partes está, a su vez, dividida en secciones, divisiones y capítulos. En resumen, los temas tratados por este manual son[21]:

TOMO I

Parte Primera: *Filosofía de la ciencia estadística, y teoría de sus principios*

SECCIÓN PRIMERA: *De la Estadística en general*
SECCIÓN SEGUNDA: *Principios filosóficos de la ciencia estadística*

Segunda Parte: *Aplicación de los principios de la ciencia estadística a la práctica de sus investigaciones*

SECCIÓN PRIMERA: *Topografía*
SECCIÓN SEGUNDA: *Investigaciones estadísticas sobre la población*
SECCIÓN TERCERA: *Producciones naturales*
SECCIÓN CUARTA: *Industria y artes*

TOMO II

SECCIÓN QUINTA: *Comercio y navegación*
SECCIÓN SEXTA: *Administración pública*

[21] Para un más detallado contenido de la obra , véase MARTÍN PLIEGO, F.J. : *Estudio Preliminar* en IBÁÑEZ, J.M. : Tratado Elemental de Estadística. Edición facsimilar. I.N.E.. Madrid. 2006.

SECCIÓN SÉPTIMA: *Usos y costumbres*
SECCIÓN OCTAVA Y ÚLTIMA: *Sobre la aplicación de las reglas y principios de la Estadística*

Por último, Ibáñez explica el cómo cumplimentar los modelos de tablas que figuran como anexo a su obra.

Discurso que en la solemne apertura de la cátedra de ciencia estadística, establecida por la Sociedad económica matritense dijo D. José María Ibáñez, del consejo de S.M., ministro honorario del suprimido tribunal supremo de hacienda, fiscal general cesante de los correos y caminos, y de la suprema junta de apelaciones de los mismos, vocal secretario de la comision de estadística de la riqueza pública, etc, y profesor elegido para dicha cátedra.

En este ampuloso y extenso discurso de corte decimonónico J.M. Ibáñez resume los argumentos que expone en su Tratado sobre la necesidad de fomentar el estudio de la ciencia de la Estadística tan necesario en las sociedades más avanzadas para poder poner de manifiesto las fuerzas y el poder de cada país. Ibáñez no logra comprender cómo siendo España un país donde se han reiterado a lo largo de su historia más que en ningún otro todo tipo de trabajos estadísticos, tales como los censos y catastros entonces conocidos, no exista sensibilidad suficiente para dar sentido conjunto a estas investigaciones estableciendo métodos adecuados para mejorar los resultados y coordinar todos los esfuerzos que se dirijan a este fin. En este sentido dijo[22]:

"*Y así como España fue la primera en formar compilaciones y relaciones estadísticas mucho antes que otras naciones tratasen este punto, no puede decirse otro tanto con respecto al estudio de una ciencia que varias de ellas han fomentado y adelantado ya considerablemente.*

Alemania es sin duda el primer país en que apareció la Estadística reducida á un sistema determinado y fijo en la aplicación de sus reglas, y sujeta a un estudio uniforme y regular. Un célebre profesor, que la dió el nombre que hoy conserva, explicó públicamente los elementos de la misma ciencia, si bien con arreglo a la opinión particular que de ella había formado considerandola casi exclusivamente como una ciencia destinada a dirijir a los hombres del Estado. Desde allí se comunicó su estudio a Inglaterra en donde hizo todavia adelantos de bastante consideración, y se extendió sucesivamente á Italia, á Francia y otros países..."

En varios pasajes de su discurso intuye la utilidad de la Estadística para establecer leyes más generales a partir de las observaciones efectuadas, claro está no encontraremos ningún atisbo de lo que devendría en estadística inferencial pero sus ejemplos y posibles aplicaciones de esta ciencia apuntan en esa dirección. Citaremos dos párrafos de su discurso al referirse a la utilidad de esta ciencia y de sus resultados. En el primero de ellos[23] dice:

"*Ella inquiere lo que existe y de que modo existe para deducir por medio de comparaciones, de proporciones y de cálculos fundados en la mas posible exactitud si*

[22] Véase "El Amigo del Pais". Tomo II, nº 20. 15 de diciembre de 1844. pág. 196. Discurso leído el 1 de diciembre de 1844.
[23] Ibidem, pág. 198

lo que existe puede recibir mejoras, aumentos, beneficios; y en que forma y bajo que sistema será mas fácil conseguirlo".

Y en el segundo[24]:

"Ellos manifiestan igualmente la relacion que hay entre la ignorancia y los delitos, y entre la instruccion y la moralidad de los individuos".

Sobre la creación de un Instituto central de Estadística de España

En este caso[25] se trata de un breve documento que realmente no contiene un proyecto completo para la creación de un Instituto de Estadística en España sino que lo que defiende es una propuesta sobre la idoneidad de su creación. Cabría esperar que bajo el título de este trabajo nos encontráramos con un plan completo para la constitución de este organismo, con un organigrama de funcionamiento y con un plan estadístico que debería desarrollar. Pero no es así, realmente su plan estadístico está contenido en su Tratado y lo que aquí leemos no es otra cosa que el resultado de su decepcionante experiencia en su paso por las comisiones estadísticas en las que participó. En este sentido deja reflejado su pesar al decir, sobre la Comisión creada en 1843 de la que fue secretario, que[26]:

" *y que según el tenor del mismo R. Decreto debía tener después su existencia legal a favor de una ley que en el mismo se prescribía presentar a los Cuerpos Colegisladores, no tuvo al fin este exito, ni la misma Comisión, de que tuve la honra de formar parte, logró que se le dieran los auxilios indispensables para emprender sus trabajos en las provincias, habiendo sido disuelta por último en Real orden de 14 de junio del presente año, cabalmente al plantearse el nuevo sistema de tributos".*

Convencido que la Administración no estaba por la labor de tomarse en serio esta cuestión, propone que sea la sociedad civil quien cree una institución que se ocupe de todos o de algunos ramos de la Estadística y al igual de lo que pasa en Francia se constituya una sociedad[27] , "*que podría denominarse, por ejemplo, Instituto central de Estadística de España"*, sirviéndole de corresponsales todas las sociedades económicas del reino y las demás corporaciones científicas y literarias. Este Instituto debería formarse tomando por base a nuestra Sociedad.

Otros escritos relativos al comercio exterior de nuestro país completan el conjunto de aportaciones de José María Ibáñez Ramos, aunque no se le escapará a nadie que el más importante es su Tratado Elemental de Estadística.

Bibliografía

ARCHIVO HISTÓRICO NACIONAL: Fondos contemporáneos, Ministerio de Justicia, Magistrados y Jueces, legajo 4496-1, expediente 3909

[24] Ibidem, pág. 199
[25] Sociedad Económica Matritense: Legajo 386, expediente 1, de fecha 1 de agosto de1845.
[26] Ibidem, pág. 2
[27] Ibidem, pag.3

ARCHIVO HISTÓRICO NACIONAL: Fondos contemporáneos, Ministerio de Hacienda, serie general, legajo 5093, expediente 73.

ARCHIVO HISTÓRICO NACIONAL: Fondos contemporáneos, Consejos, legajo 12102, expediente 152.

CARRERAS Y GONZALEZ, M; PIERNAS Y HURTADO, J.M. : *Tratado Elemental de Estadística*. Imprenta y Librería de Miguel Guijarro. Madrid, 1873.

DALMAU, F. : *Ensayos de Estadística practicados en la provincia de Granada*. Imprenta de Ibarra. Madrid. 1820.

DUFAU, P.A. : *Traité de statistique, ou Théorie de l´étude des lois d´après lesquelles se développent les faits sociaux, suivi d´un essai de statistique physique et morale de la population française*. Delloye. Paris. 1840.

DUFAU, P.A.: *Tratado de Estadística o Teoría del estudio de las leyes según las cuales se desarrollan los hechos sociales seguido de un ensayo de estadística física y moral de la población francesa.* Imprenta y Librería de Don Ignacio Boix. Madrid. 1845.

GIOJA, M. : *Filosofia della Statistica*. Presso Giovanni Pirotta. Milano. 1826.

HERRERA DÁVILA, J. ; ALVEAR, A. : *Lecciones de Estadística*. Imprenta de D. Mariano Caro. Sevilla. 1829.

LABORDA, A.: *Textos de Estadística. José María Ibáñez*. La hoja del Monte. Madrid. 2006.

IBÁÑEZ, J.M.: *Tratado Elemental de Estadística*. Imprenta del Colegio de Sordo-mudos. Madrid. Tomo I, 1844. Tomo II, 1845.

IBÁÑEZ, J.M.: *Tratado Elemental de Estadística .Tomos I y II*. I.N.E.. Madrid.2006.

MADOZ, P: *Diccionario Geográfico-Estadístico-Histórico de España y sus posesiones de ultramar*. Imprenta de D. Pascual Madoz. Madrid. 1848.

MARTÍN PLIEGO, F.J. : "Estudio Preliminar en IBÁÑEZ, J.M." : *Tratado Elemental de Estadística*. Edición facsimilar. I.N.E.. Madrid. 2006.

MOREAU DE JONNES, A. : *Elementos de Estadística. Principios generales de esta ciencia, su clasificación, metodo, operaciones, diversos grados de certidumbre, errores y progresos, con su aplicación a la comprobación de los hechos naturales, sociales y políticos, históricos y contemporáneos*. Imprenta de Francisco Abienzo. Madrid. 1857.

PEUCHET, J. : *Essai d´une Statistique Génerale de la France. Chez Testu*, imprimeur-libraire. Paris. 1798.

PIERNAS HURTADO. J. : *Tratado Elemental de Estadística*. Librería de Victoriano Suárez. Madrid. 1897.

POU Y ORDINAS, A. J. : *Curso de Estadística.* Imprenta de la Viuda e H. de J. Subirana. Barcelona. 1889.

SALVÁ, M.: *Tratado Elemental de Estadística. Agustín Jubera*. Madrid. 1881.

SANZ SERRANO; A: *Resumen Histórico de la Estadística en España*. INE.

Madrid. 1956

SÁNCHEZ-LAFUENTE FERNÁNDEZ, J. : *Historia de la Estadística como ciencia en España* (1500-1900). INE. Madrid. 1975.

SOCIEDAD ECONÓMICA MATRITENSE : Legajo 386 expediente 1. 1845.

SOCIEDAD ECONÓMICA MATRITENSE : El Amigo del Pais, Tomo II nº 20. 1844.

Velasco Pérez, R. : " Una nota sobre José María Ibáñez, primer catedrático de Estadística en España" en Homenaje al Profesor Juan Sánchez-Lafuente. Facultad de Ciencias Económicas y Empresariales de la Universidad de Málaga. Málaga. 1990.

Capítulo 4

La memoria de Daniel Bernoulli sobre la inoculación contra la viruela (1760): Un problema de decisión bajo incertidumbre

JOSÉ ANTONIO CAMÚÑEZ RUIZ
JESÚS BASULTO SANTOS
FRANCISCO JAVIER ORTEGA IRIZO
Universidad de Sevilla

Sobre la inoculación y la posterior vacunación como lucha contra la viruela

La viruela (en francés, *petite vérole*, en inglés, *smallpox*) es una enfermedad infecciosa de origen viral, muy contagiosa, y que ha sido totalmente erradicada a partir de 1977 (fecha del último caso conocido, que apareció en Somalia). Pero, la viruela fue una plaga terrible y temida. Los cálculos situaban su mortalidad en una tasa de un enfermo de cada ocho o incluso mayor, según el momento del análisis. La enfermedad siempre quedó fuera del alcance de un tratamiento eficaz, siendo el único remedio contra ella el de la vacunación.

Previo a la vacunación se usó una técnica conocida como “inoculación” o “variolización”, con mayor riesgo que la vacunación, y que tuvo su auge en Europa durante el siglo XVIII, dando lugar a una importante serie de investigaciones y discusiones por parte de médicos, matemáticos, estadísticos y, en general, pensadores, a lo largo de dicho siglo, hasta la aparición de la vacuna. La nómina de investigadores sobre el asunto es extensa para ese periodo. Así, podemos citar, además de la Memoria de Daniel Bernoulli, las investigaciones de Charles Marie de La Condamine (1701-1774), Jean le Rond D’Alembert (1717-1783), Johann Heinrich Lambert (1728-1777), Jean Trembley (1749-1811), Emmanuel Étienne Duvillard (1755-1832) y Joshua Milne (1776-1851).

Asociado a la viruela aparece el término “variola”, palabra derivada del latín “varius” que significa grano o pústula. El contagio de la enfermedad es directo, de humano a humano, pues no existen reservorios animales. La entrada del virus se produce normalmente a través de las vías respiratorias. El periodo de incubación es de 10 a 14 días. Aparecen escalofríos, fiebres, dolor de cabeza, náuseas… Tras estos síntomas se produce la erupción, caracterizada por la aparición de manchas rojas sobre la piel, que se convierten en vesículas, después en pústulas llenas de pus, dolorosas, densas y redondas, para acabar formando costras. Estas manifestaciones aparecen primero en la mucosa de la boca y faringe, luego en la cara y extremidades, para pasar después al tronco, a las palmas de las manos y pies. Las costras aparecen cerca del octavo o noveno día de evolución y, cuando se desprenden, dejan una cicatriz en la piel. Por tanto, es una enfermedad deformante que deja sus huellas para toda la vida. Las marcas de viruela se observan en el 70% de los supervivientes siendo las lesiones en el rostro las que prevalecen por la tendencia a la infección de las glándulas sebáceas. La enfermedad, si no mata al paciente, es inmunizante: cualquier infección por el mismo virus es imposible durante años.

Como primera práctica para erradicar el mal se usó la "inoculación", que consistía en introducir un hilo de seda impregnado de pus tomada de una pústula de viruela de alguien que sufriese la enfermedad, en una ligera incisión hecha en el brazo de la persona a inocular. El hilo era retirado al cabo de dos días. Se preparaba al paciente con una dieta ligera y se le aislaba durante una semana justo hasta que un breve acceso de fiebre junto con una erupción local marcaba la aparición de la enfermedad "artificial". El procedimiento era arriesgado, pues los inoculados podían acabar falleciendo por la propia enfermedad, contraída en este caso de forma voluntaria. Había estimaciones diversas, nada coincidentes, de la tasa de ese riesgo.

La técnica era muy antigua en China y se fue difundiendo a lo largo de la ruta de la seda hacia occidente. En 1701, Giacomo Pylarini realiza la primera inoculación en Constantinopla. Llega a Gran Bretaña a través de Lady Mary Wortlley Montagu (1689-1762), esposa del embajador de este país en Turquía, mujer de alta sociedad y de una fuerte formación intelectual, escritora e independiente, que aprende la técnica del doctor Emmanuel Timoni (1670-1718). El doctor Timoni, médico de la embajada en Estambul, diplomado por la Universidad de Padua, y miembro de la Royal Society de Londres, publica en 1713 en las Philosophical Transactions de esta sociedad su tratado sobre la inoculación. Pero la concienciación de la sociedad británica sobre la inoculación se debe a Lady Mary Wortlley Montagu, quien inoculó a sus hijos y aprovechó su influencia y sus escritos para llegar a la misma Reina que hizo lo propio con los suyos.

La moda de inocular se propagó por Francia a partir de 1755, después de que M. Hosty, doctor de la facultad de París, fuese instruido en Londres con el método.

La discusión sobre la eficacia de la inoculación en la lucha contra la viruela se mantuvo en los círculos científicos europeos durante la segunda mitad del siglo XVIII, hasta la aparición de las investigaciones de Edward Jenner (1749-1823), médico y naturalista inglés, el cuál, desde que empieza a ejercer como médico, se muestra partidario de la inoculación y la lleva a la práctica. Pero recoge de sus pacientes la existencia de una extraña creencia popular: las personas que trabajan en las granjas de la región nunca son alcanzadas por la viruela.

Esto le lleva a observar un hecho: los animales de granja, vacas, equinos y cerdos sufren enfermedades parecidas a la viruela, conocidas como Cowpox, Horsepox y Swinepox, y dichas enfermedades pueden transmitirse al hombre. Confirmó que las granjeras que ordeñaban las vacas que sufrían de Cowpox, desarrollaban unas pocas pústulas en sus manos y, además, se mantenían protegidas contra la viruela, porque no enfermaban en un brote epidémico.

Entre 1789 y 1796 realiza diversos experimentos de inoculación con smallpox y cowpox concluyendo con una prueba final en 1796, cuando varioliza un niño de 8 años con fluido vesicular de las pústulas de una granjera que se había contagiado de cowpox al ordeñar su vaca. A los ocho días aparecieron en la zona de la inoculación unas pocas pústulas y el niño tuvo una fiebre ligera. Pero el gran desafío se produjo dos meses después. Jenner inocula al niño una fuerte dosis de viruela, tan fuerte que debería enfermarlo. Sin embargo, nada ocurre, el cowpox protegió contra la viruela. ¡Había nacido la vacuna! Nombre que fue adoptado por el origen del material obtenido a partir de la vaca (Vaccinus, vacca en latín). Un siglo después, Pasteur extiende el nombre de vacunación a la inmunización contra otros agentes.

Después de repetir el experimento en varios sujetos más, Jenner escribe un trabajo para su publicación en "Philosophical Transaction", pero fue rechazado. Por tanto, para dar publicidad a su descubrimiento, este autor escribe un pequeño texto cuya impresión debe financiar él mismo. La importancia de la experiencia de este científico fue reconocida inmediatamente y su libro fue traducido a seis idiomas de modo que la vacuna comenzó a emplearse fuera de Inglaterra en muchos países de Europa.

Ni que decir tiene, la forma de conservar la vacuna, pasando de brazo a brazo, era compleja, resultando en una cadena viviente no siempre fácil de controlar. Era más cómodo desplazarse con un hombre recién vacunado y que presentase pústulas, que con una vaca infectada. Con el tiempo, este sistema trajo problemas de contaminación (Louis Pasteur y la asepsia no llegan hasta más tarde). De todas formas, la base fundamental de la vacunación quedó establecida a partir del trabajo de Jenner.

La memoria de Daniel Bernoulli (1700-1782)

Este autor, físico y matemático, que ejerció cátedra de botánica, de anatomía, de filosofía natural y de matemáticas, era miembro de una familia suiza que dio al mundo ocho notables matemáticos en tres generaciones.

El 30 de abril de 1760, D. Bernolli leyó una memoria en la Real Academia de las Ciencias de París titulada *Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir*. La memoria de D. Bernoulli, sin embargo, no fue publicada hasta 1765. Esto no fue, seguramente, por azar teniendo en cuenta que d'Alembert, que ocupaba en la Academia una posición influyente, consiguió hacer publicar la suya en 1761. Este último criticó la solución del problema de la inoculación dado por "un sabio geómetra", sin nombrarlo. Hemos de reconocer, sin embargo, que otras memorias leídas en la Academia en ese período también sufrieron retraso en su publicación. Por ejemplo, la de Trembley fue leída en 1796 y publicada en 1799.

D'Alembert no coge el punto de vista de la colectividad, sino únicamente el del individuo que debe elegir entre el riesgo inmediato de la inoculación, y el riesgo más distante de la enfermedad: se goza mejor de la vida cuando se es joven, dice él, y una ganancia de tres o cuatro años de vida media, perspectiva lejana (¿años de vejez?) no es suficiente para justificar que se exponga a morir en pocos días de una inoculación voluntaria.

D. Bernoulli, añadió una "Introducción Apologética" para la publicación de 1765. En ella declara que no estaba *sorprendido de que al vulgo le llame poco la atención este último aspecto, pero no puedo impedir estarlo cuando veo personas de mérito y de una gran reputación, plantearse seriamente si vale la pena sufrir una operación como la inoculación, con la esperanza de prolongar su vida en dos años: sería de desear que las críticas fuesen más reservadas y más circunspectas, y sobretodo que hiciesen el esfuerzo de ponerse en el hecho de las cosas que se proponen criticar por anticipado.* Como vemos, la nota es bastante mordaz, dada la cortesía de la época. La debilidad del texto de d'Alembert viene del hecho de su resistencia a la teoría de las probabilidades. Trembley (1799) lo dice sin ambages: "ha sustituido el análisis elegante del señor Bernoulli por una teoría matemática tan fuertemente matemática, que ni él ni nadie, que yo sepa, ha hecho aplicación". "El ilustre d'Alembert se precipita en la contienda", dice Duvillard (1806), apuntando algunos errores en la metafísica del cálculo que extrañan en el caso de "un tan grande geómetra", y añade una alabanza al trabajo de Bernoulli: "Este escrito, el mejor sin duda que haya sido publicado a favor de la inoculación, es notable por la finura, la precisión de las ideas, las visiones y las cosas que encierra" .

Los argumentos de d'Alembert tienen sin embargo el interés de atraer la atención sobre dos hechos: las estimaciones relativas a la esperanza de vida de una población no son más que un elemento de valoración, entre otros, cuando se razona sobre la chance de sobrevivir de un individuo dotado de una cierta constitución; y la utilidad de una determinada solución debe evaluarse no sólo con la *cantidad* de vida que ofrece, sino también con su *calidad*.

Volviendo al trabajo de Bernoulli, éste dice haber compuesto su memoria sobre la mortalidad de la viruela a petición de Maupertuis (1698-1759, matemático, astrónomo y biólogo francés, con importantes trabajos en genética), que le había confiado su proyecto de exponer *"en una misma Tabla los dos estados de la humanidad, uno tal como es efectivamente, y otro tal como sería si se pudiese eximir de la viruela todo el género humano"*. Así, en esta importante memoria, D. Bernoulli presentó la primera doble tabla de vida decreciente de la historia de la ciencia, además de proponer un modelo matemático de comportamiento de la viruela en una población. Y el autor añade: *"Pienso que el paralelismo de esto dos estados explicaría mejor la diferencia y el contraste, que cualquier comentario más amplio, pero he sentido también la dificultad del intento; y la insuficiencia de las listas mortuorias, que no señalan la edad de aquellos que la viruela se lleva; no hace más que poner un mayor obstáculo a estos enfoques."*

Así pues, la gran debilidad de este trabajo viene de la falta de datos empíricos. El único dato contrastado con el que D. Bernoulli contaba (según las listas de Süssmich, 1741, entre otros) es el de la proporción de mortalidad de viruela sobre el total de muertes, considerando todas las edades: *"Está constatado por una larga serie de observaciones que la viruela se lleva la treceava o catorceava parte de cada generación"*. O sea, parte del supuesto contrastado que la suma de todas las muertes de viruela hace alrededor de 1/13 del total de las muertes, o sea, 100 sobre su generación de 1300 (de hecho, su tabla suma 101). En cuanto a lo que se llamará más tarde la letalidad de la enfermedad, este autor no tiene más que estimaciones muy fragmentarias: *"Se sabe… que esta enfermedad se lleva alrededor de la octava o la séptima parte de aquellos a los que le ataca, con tal que se tome la proporción sobre un gran número de epidemias; …"* Esto le servirá de soporte sobre el que apoyar las proporciones que introducirá después.

D. Bernoulli toma como base de cálculo la tabla de mortalidad establecida por el astrónomo inglés Halley (1693) que preparó con datos de la ciudad de Breslaw, y que nuestro autor conoció a través del autor alemán Süssmilch. Esta tabla no indica el número de niños recién nacidos. Su punto de partida es el número de supervivientes a la edad de un año, que considera igual a mil. Süssmilch (1741) supone que Halley había partido de una generación de 1238 niños, número medio de nacimientos en Breslaw para el período considerado. D. Bernoulli no cree admisible este supuesto, porque las tablas de Buffon, publicadas en 1749, daban una mortalidad infantil mucho más elevada. Entonces, toma como punto de partida a la edad de cero años una cifra arbitraria de 1300 niños, intermedia entre las estimaciones de Buffon y de Süssmilch.

Falto de datos, D. Bernoulli se limita a efectuar conjeturas. Apela a los párrocos de las ciudades para que anoten las edades de los fallecidos (p. 22), a los médicos para que valoren la importancia de la inoculación (p. 29), y afirma que el día en que se tenga listas mortuorias bien hechas, él será incluso su "más severo crítico" (p. 4). Su método de cálculo no es, desgraciadamente, aplicable bajo hipótesis distintas a las suyas, como lo señala Trembley (1799). Pero ve claramente lo que está en juego en el problema; un conocimiento exacto de los riesgos permitiría no sólo a los particulares, sino también a los Estados, tomar decisiones racionales: *Si se conociesen con exactitud todas las proporciones medias que se hubiesen podido determinar sobre un número muy grande de observaciones, pero bien consideradas y reflexionadas, se podría dar una teoría completa sobre los azares de la viruela: una teoría así dictaría las máximas que todo hombre razonable debe seguir* (p. 8).

Los cálculos de D. Bernoulli

El problema planteado es encontrar una fórmula que permita obtener, a partir de la tabla de mortalidad en "el estado natural y con viruela", una tabla de mortalidad del "estado sin viruela".

Para llevar a cabo su propósito, el autor echa en falta dos conocimientos: *"Veía mejor, en primer lugar, que la realización de una idea así exige dos conocimientos elementales: ¿cuál es el riesgo anual a diferentes edades de ser sorprendido por la viruela?, ¿y cuál es el de morir para aquellos que son atacados?"*

Entonces, D. Bernoulli plantea dos principios: el primero es que "tanto que no se haya tenido la viruela, se corre continuamente el mismo riesgo de tenerla", el segundo es "que el riesgo de morir de viruela cuando se es atacado, (es)… el mismo en toda edad" (1760b, Introducción, p. 4).

En la Introducción que, como se ha dicho fue, redactada en el momento de la publicación, en 1765, da algunos argumentos que intentan mostrar que las hipótesis elegidas son, al menos, verosímiles. El texto primitivo de la memoria insiste en el hecho de que se puede adaptar el cálculo a las condiciones locales: así, en el "pequeño país" de Bernoulli (Bâle), los médicos estimarían 1/20 la probabilidad de morir cuando se contrae la viruela. Pero las observaciones recogidas en las "grandes ciudades" dan una probabilidad mucho mayor.

Por tanto, el autor se ve en la obligación de introducir supuestos o hipótesis con los que poder construir sus cálculos. Asume que nadie puede padecer viruela más de una vez en la vida, e introduce dos proporciones que considera constantes para las diferentes edades: la proporción anual de los que cogen viruela entre aquellos que nunca la han tenido es $1/n$, o sea, de cada n individuos que no hayan tenido viruela uno la coge en el plazo de un año, y que la proporción anual de muerte entre aquellos que enferman de viruela es $1/m$, o sea, de cada m individuos que enferman de viruela en el mismo año, uno de ellos fallece de esa enfermedad. Por tanto, según esto, la probabilidad de que un individuo que hasta ahora no ha padecido de viruela, muera de esa enfermedad en el plazo de un año será el producto de las dos anteriores, esto es $\frac{1}{nm}$. Escribe argumentos para mantener su hipótesis de que se puede considerar la misma proporción de aquellos que cogen la viruela en todos los tramos de edad. Concluye diciendo: *En tanto no se haya tenido la viruela se corre continuamente el mismo riesgo de tenerla. Las leyes de la Naturaleza más simples son las más verosímiles.*

Con este planteamiento, el problema es qué valores hay que dar a los dos parámetros *n* y *m*, ante lo ya comentado de ausencia de datos de enfermos y fallecidos de viruela por edades. Este asunto será motivo de discusión y de análisis entre los investigadores posteriores a Bernoulli en el siglo XVIII. Cada uno de ellos proporcionará valores distintos para *n* y *m* según los tramos de edad y en base a algunas estadísticas de las que ya se disponían.

Bernoulli propone como valores de ambos parámetros $n = m = 8$. O sea, la probabilidad de contraer la enfermedad, en el plazo de un año, para un individuo que no la ha tenido es 1/8, para cualquiera que sea su edad, y la probabilidad de fallecer de viruela en dicho plazo es 1/64. El autor elige esos valores en base a los escasos datos de que disponía y, además, porque no contradecía al dato más sustentado de una muerte de cada 13 era debida a la viruela.

Con estas premisas se propone construir una fórmula que, bajo supuestos razonables, proporcione el número "*s*" de personas que no han tenido la viruela, de una edad "*x*", en función de dicha edad y del número "*y*" de supervivientes. Para ello, argumenta como sigue. Los supervivientes, *s*, que no han tenido viruela decrecen por

(i) aquellos que cogen viruela (muriendo o no de ello) y,
(ii) aquellos que mueren de otras causas sin haber tenido viruela alguna vez.

En un elemento de tiempo dx el decrecimiento de *s* es $-ds$ (ds y dy, usado abajo, son intrínsecamente negativos y el signo menos es entonces necesario para convertirlos en números positivos). El número de atacados por viruela es $\frac{sdx}{n}$, y el número de aquellos que mueren de viruela es $\frac{sdx}{nm}$. El número total de muertes por todas las causas en un tiempo dx es $-dy$. Por tanto, el número de muertes por otras causas distintas a la viruela es $-dy - \frac{sdx}{nm}$. Pero este número relaciona a *y* personas, mientras que al formar una ecuación para *s* hemos referido el número de muertes de otras causas entre *s*, esto es, con $\frac{\left(-dy - \frac{sdx}{nm}\right)s}{y}$ muertes.

Entonces

$$-ds = \frac{sdx}{n} + \frac{\left(-dy - \frac{sdx}{nm}\right)s}{y}.$$

Por tanto

$$\frac{s \cdot dy - y \cdot ds}{s^2} = \frac{ydx}{sn} - \frac{dx}{nm}.$$

Poniendo $\frac{y}{s}=r$, por lo que $dr=\frac{s\cdot dy-y\cdot ds}{s^2}$. Entonces, $nm\cdot dr=mr\cdot dx-dx$, o $dx=\frac{nm\cdot dr}{mr-1}$. Integrando da $n\log(mr-1)=x+c$, donde c es una constante a determinar, o $n\log\left(\frac{my}{s}-1\right)=x+c$. Ahora, cuando $x=0$, $y=s$, lo que da

$$c=n\log(m-1)$$

Y de aquí

$$n\log\left(\frac{\frac{my}{s}-1}{m-1}\right)=x.$$

Entonces

$$\frac{my}{s}-1=(m-1)e^{x/n}.$$

Por tanto

$$s=\frac{m}{1+(m-1)e^{x/n}}y. \qquad (1)$$

Para los valores que estableció para *m* y *n*, la fórmula queda

$$s=\frac{8}{7e^{\frac{x}{8}}+1}y.$$

Con esto, el autor construye la Tabla 1 de su memoria, donde en la primera columna aparece la edad (lo que hemos llamado *x*); en la segunda, el número de supervivientes a esa edad según la tabla de Halley (lo que hemos llamado *y*).

Tabla 1.

Edad por años	Supervivientes según Halley	No habiendo tenido la viruela	Habiendo tenido la viruela	Cogiendo la viruela durante cada año	Muertes de viruela cada año	Suma de muertes de viruela	Muertes de otras enfermedades cada año
0	1300	1300	0				
1	1000	896	104	137	17’1	17’1	283
2	855	685	170	99	12’4	29’5	133
3	798	571	227	78	9’7	39’2	47
4	760	485	275	66	8’3	47’5	30
5	732	416	316	56	7’0	54’5	21
6	710	359	351	48	6’0	60’5	16
7	692	311	381	42	5’2	65’7	12’8
8	680	272	408	36	4’5	70’2	7’5
9	670	237	433	32	4’0	74’2	6
10	661	208	453	28	3’5	77’7	5’5
11	653	182	471	24’4	3’0	80’7	5
12	646	160	486	21’4	2’7	83’4	4’3
13	640	140	500	18’7	2’3	85’7	3’7
14	634	123	511	16’6	2’1	87’8	3’9
15	628	108	520	14’4	1’8	89’6	4’2
16	622	94	528	12’6	1’6	91’2	4’4
17	616	83	533	11’0	1’4	92’6	4’6
18	610	72	538	9’7	1’2	93’8	4’8
19	604	63	541	8’4	1’0	94’8	5
20	598	56	542	7’4	0’9	95’7	5’1
21	592	48’5	543	6’5	0’8	96’5	5’2
22	586	42’5	543	5’6	0’7	97’2	5’3
23	579	37	542	5’0	0’6	97’8	6’4
24	572	32’4	540	4’4	0’5	98’3	6’5

En la tercera columna, nuestro presenta el número de personas que han llegado a esa edad sin haber padecido aún la viruela (lo que hemos llamado *s*), habiendo sido calculado ese número mediante la fórmula anterior última; en la cuarta, el número de individuos que, a esa edad, ya han padecido viruela y se han recuperado (supervivientes de la enfermedad), o sea $y-s$; en la quinta, el número de enfermos de viruela de ese año, por tanto, $\frac{1}{8}s$, pero para mayor exactitud toma para *s* la media entre los valores de esa cantidad para ese año y el anterior, o sea, construye la 5ª columna calculando $\frac{1}{8}\times\frac{s_{x-1}+s_x}{2}$; en la sexta, el número de fallecidos de viruela ese año: un octavo de los valores de la quinta columna; en la séptima, las muertes acumuladas de viruela desde 0 años hasta ese año y, por último, en la octava, las muertes por otras causas distintas a la viruela, o sea, $y_{x-1} - y_x - \text{columna } 6$. La Tabla se construye hasta la edad de 24 años porque, como se ha dicho, para esa edad, pocos de los que sobreviven no han padecido la enfermedad.

Resulta de las tablas calculadas por D. Bernoulli que en el estado natural, para una generación de 1300 recién nacidos, 500 morirán sin haber contraído la viruela, 700 la cogerán y morirán de otra cosa y 100 cogerán la viruela y morirán por culpa de ella. La mitad de estos últimos, mueren antes de los cinco años.

A continuación, nuestro autor considera lo que ocurriría si todos fuesen inoculados al nacer y, entonces, la viruela fuese erradicada como causa de muerte. Eso le permite preparar una tabla de vida para el caso en que no hubiese muerte por viruela. Ilustramos el método de construcción de esta segunda tabla usando los números de la Tabla 1:

(iii) En el primer año de vida hay 17’1 muertes de viruela, así que, sin viruela, el número de los que sobreviven un año crecerá de 1000 a 1017’1.

(iv) Si 133 (columna 8) mueren durante el segundo año de otras causas distintas a la viruela entre los 1000 vivos al comienzo de este año, sería, por proporción, 135’3 muertes entre los 1017’1 vivos al comienzo de ese año en una situación libre de viruela, quedando 881’8 vivos al final del segundo año; y así para el resto de edades.

Los resultados son recogidos en la Tabla 2, donde se comparan las columnas de supervivientes en el estado natural (esto es, con muertos de viruela) y en el estado no varioloso.

Al construir esta segunda tabla, Bernoulli descuida el hecho de que algunos de estos supervivientes podrían sucumbir en ese año de otras enfermedades, algo que Duvillard (1806) rectificará.

Tabla 2.

Edad por Años	Estado natural y varioloso	Estado no varioloso	Diferencia o ganancia	Edad por Años	Estado natural y varioloso	Estado no varioloso	Diferencia o ganancia
0	1300	1300	0	13	640	741’1	74’1
1	1000	1017’1	17’1	14	634	709’7	75’7
2	855	881’8	26’8	15	628	705’0	77’0
3	798	833’3	35’3	16	622	700’1	78’1
4	760	802’0	42’0	17	616	695’0	79’0
5	732	779’8	47’8	18	610	689’6	79’6
6	710	762’8	52’8	19	604	684’0	80’0
7	692	749’1	57’2	20	598	678’2	80’2
8	680	740’9	60’9	21	592	672’3	80’3
9	670	734’4	64’4	22	586	666’3	80’3
10	661	728’4	67’4	23	579	659’0	80’0
11	653	722’9	69’9	24	572	651’7	79’7
12	646	718’2	72’2	25	565	644’3	79’3

Esta Tabla hace ver de un vistazo, cuántos de 1300 niños, supuestos nacidos al mismo tiempo, quedarían vivos de año en año hasta la edad de veinticinco años, suponiéndoles todos sujetos a la viruela; y cuántos quedarían si estuviesen todos exentos de esta enfermedad, con la comparación y diferencia de los dos estados.

Sin emplear la expresión “esperanza de vida”, Daniel Bernoulli calcula que la vida media para un recién nacido, en el estado no varioloso, sería de 29 años y 9 meses (contra 26 años y 7 meses en el estado varioloso), o sea, una ganancia de 3 años y 1 mes. A continuación plantea el riesgo de la inoculación. Inicialmente, plantea dicho riesgo como de 1 sobre *N*. Por lo tanto, la generación inicial ha de ser disminuida en la proporción de N a $N-1$. Comenta que no hay acuerdo en cuánto debe ser *N*. Opina que en el peor de los casos $N=200$ y, entonces, bajo ese supuesto calcula que el riesgo de la inoculación supone una disminución inferior a dos meses de la vida media calculada para el estado no varioloso. Por tanto, la inoculación, con el riesgo añadido, supone una ganancia en la vida media de 3 años, sobre una media inicial de 26 años y 7 meses en el estado natural o varioloso, o sea, $\frac{1}{9}$ de esa media.

En este contexto, Bernoulli se plantea cuál debería ser *N* para que la inoculación no alterase la vida media del estado natural. El cálculo le lleva a $N=9'43$. Entonces, añade que si la inoculación se lleva menos de 100 sobre 943 hará más bien que mal a la humanidad, sobre todo si es practicada en los más jóvenes, pues entonces “la pérdida sólo caería en los niños inútiles para la sociedad, y… toda la ganancia recaería sobre esta edad que es la más preciosa” (edad adulta). D. Bernoulli se anticipa aquí a los estudios de Quételet, tendientes a medir el coste, para la sociedad, de la mortalidad de los jóvenes clasificados por edad.

Sólo después de haber preparado su Tabla 2, Bernoulli se da cuenta que era posible obtener una fórmula que relacionase los supervivientes de las dos tablas. Como se demuestra arriba las muertes en el período de tiempo dx por causas distintas a la viruela es $-dy-\dfrac{sdx}{nm}$ con respecto de una población de y (supervivientes en el estado natural). Por tanto, para una población de z (supervivientes en el estado no varioloso) podemos escribir

$$-dz=-\frac{z}{y}\left(dy+\frac{sdx}{nm}\right)$$

O bien,

$$\frac{dz}{z}-\frac{dy}{y}=\frac{sdx}{nmy}.$$

Sustituyendo (1) para un s dado

$$\frac{dz}{z}-\frac{dy}{y}=\frac{1}{n}\left\{1-\frac{(m-1)e^{\frac{x}{n}}}{1+(m-1)e^{\frac{x}{n}}}\right\}dx.$$

Integrando obtenemos

$$\log\frac{z}{y}=\frac{x}{n}-\log\left\{1+(m-1)e^{\frac{x}{n}}\right\}dx+c.$$

Dado que $y=z$ cuando $x=0$, entonces $c=\log m$, de donde se deduce

$$z=\frac{me^{\frac{x}{n}}}{1+(m-1)e^{\frac{x}{n}}}y. \qquad (2)$$

Bernoulli da sólo dos comparaciones numéricas de z por su método aproximado para construir la Tabla 2 y por la fórmula exacta (2); estos son:

Valores de z

	Aproximado	Exacto
Edad (x)	(Tabla 2)	(Fórmula 2)
16	700’1	697’4
24	651’7	649’2

Tras considerar los dos ejemplos comenta el acuerdo razonable entre su método aproximado y su fórmula exacta. En cada caso el valor exacto es menor que el aproximado, lo cual es de esperar ya que, como se ha dicho, el principal defecto del método aproximado de Bernoulli es que no tiene en cuenta las muertes por otras causas que tendrían lugar entre los muertos de viruela salvados en cada año en particular. Desde un punto de vista más matemático podemos entender la Tabla 2 como una forma discreta de mostrar la supervivencia, mientras que la Fórmula 2 nos da una aproximación continua del mismo asunto, introduciendo la continuidad su corrección hacia abajo.

Un problema de decisión

Aunque no disimula que él cree personalmente que la inoculación es "muy útil", D. Bernoulli declara que su intención es únicamente arrojar sobre la cuestión "alguna luz", con el fin de que se juzgue "con todo el conocimiento de causa posible".

Concibe que el problema se plantea de forma diferente a los particulares y a lo Estados. En cualquier caso, la decisión está por tomar: ¿Inoculación o no inoculación? El fundamento de esta decisión es enunciado por D. Bernoulli de manera perfectamente clara: "en tanto que se quiera adoptar *el principio de la mayor utilidad* de toda la humanidad…"

Además del cálculo de la esperanza de vida de un recién nacido en los dos estados, Bernoulli se encarga de introducir una serie de conclusiones basadas en los cálculos que realiza usando como soporte la dos tablas construidas, de manera que la toma de la decisión adecuada no deje lugar a dudas. Así, introduce el concepto de "ganancia relativa" para cada edad, cociente entre la ganancia absoluta expresada por la cuarta columna de la 2ª tabla y el número de supervivientes en el estado varioloso a esa edad. Demuestra que si todos los recién nacidos son protegidos de la viruela desde el nacimiento, la "ganancia relativa" en vidas humanas que resultaría de la erradicación de esta causa de mortalidad, crecería de año en año, llegando a 1/7 de la generación en el momento en que ésta alcance su 17º año, que Daniel Bernoulli llama su "nacimiento civil", y concluye que la proporción asintótica de "ganancia sobre los vivos" es, en general, igual a $\frac{1}{m-1}$. También, que la proporción de "vivos en estado natural"/"vivos en estado no varioloso" se estabiliza con la edad, quedando en $\frac{m-1}{m}$.

Un segundo dato que destacamos de entre la gran cantidad de cálculos que aporta el autor es: En el estado natural toda una generación queda reducida a la mitad cuando han cumplido 11 años y 5 meses, mientras que en el estado no varioloso esto ocurre a la edad de 24 años y 3 meses.

Un tercer dato es el cálculo de las esperanzas de vida de un niño de 5 años si no es inoculado y si lo es y obtiene para las mismas la proporción de 17 a 19.

El siguiente resumen en cifras es bastante contundente: En el estado natural, de 1300 individuos, 800 cogerán la viruela. Si todos son inoculados, entonces los 1300 cogerán la viruela. Por tanto, la proporción de enfermos de viruela de un estado a otro será de $\frac{8}{13}$. Pero la infección de los inoculados es mucho menor, reduciéndose en $\frac{1}{13}$. Por tanto, la proporción de enfermos de viruela de un estado a otro queda como $\frac{8}{1}$. Por otra parte, la "superficie del cuerpo de los inoculados" es 4 veces menor, por lo que, concluye, la proporción queda $\frac{32}{1}$.

La conclusión no da lugar a equívoco: aunque la inoculación sea costosa en vidas humanas, "el interés público" es practicarla, y hacerla lo más pronto posible, en cualquier caso, antes de cumplido los 5 años, pues a esa edad ya se ha llevado la mitad de aquellos que debe hacer morir. "Será siempre geométricamente cierto que el interés de los Príncipes es favorecer y proteger la inoculación…" El mejor partido es el que da mayor vida media, porque la colectividad tiene la ventaja en que los individuos acceden a una edad donde acaban siendo productivos.

Para las circunstancias particulares de un individuo, Bernoulli indica que habría que hacer cálculos particulares para conocer el partido más ventajoso. Imaginemos una persona que llega a los veinte años sin haber tenido la viruela. La sociedad no tiene casi ningún interés en hacerla vacunar a una edad donde la viruela ya se ha llevado el 95% de su tributo. La persona, por el contrario, "llegada a la edad de la razón", está en medida de sopesar la ganancia esperada de la inoculación, al riesgo que debería pagarle, y ver dónde está su ventaja. Se trata, una vez más, de elegir una solución que maximiza una esperanza matemática.

Este párrafo es bastante esclarecedor para la resolución del problema de decisión: *"Todo hombre que no ha tenido la viruela se encuentra en la agobiante necesidad de jugar durante cada año de su vida con otros 63 a cuál debe morir de esta enfermedad, y con otros 7, a cuál debe cogerla, y lleva con él esta triste suerte hasta que coge la enfermedad. ¿No es mejor, suponiendo que la inoculación quita 1 de 473, jugar contra 472 en lugar de con 63, y no tener que sufrir la suerte más que una sola vez, en lugar de que le vuelva cada año de su vida? ¿Un hombre avisado puede titubear sobre la elección? Sin embargo, esta alternativa es exactamente la de esperar la viruela natural o hacerse inocular."*

Conclusión

Estamos ante un trabajo que se nos muestra competente, científico y pragmático. El autor sabe exactamente lo que quiere hacer y procede a hacerlo. Cuando sus datos son escasos, establece supuestos y, con ellos, comprueba la coherencia con la información disponible. Usa el cálculo diferencial bajo supuestos realistas, y lleva a cabo, por primera vez en la historia de la ciencia, la comparación de las tablas de vida entre el estado natural y el estado en que quedaría al erradicarse una enfermedad.

Bibliografía

BERNOULLI, D. (1760). Essai d'une nouvelle analyse de la mortalité cause par la petite vérole, et des aventages de l'inoculation pour la prévenir, *Memoires de mathématiques et de physiques tires des registres de l'Academie Royale des Sciences, de l'année 1760; Hist. de l'Academie*. Paris, 1766, 1-45.

DAW, R. H. (1979). Smallpox and the Double Decrement Table. A piece of Actuarial Pre-History. *JIA* 106, 299-318.

DUVILLARD, E. E. (1806). *Analyse et Tableaux de l'influence de la petite vérole à chaque age, et de celle qu'un preservative tel que la vaccine peur avoir sur population et la longèvite* (París) (A. S.)

FAGOT-LARGEAULT, A. (1989). *Las causes de la mort. Histoire naturelle et facteurs de risque.* Institut Interdisciplinaire d'Etudes Epistémologiques. París.

HALLEY, E. (1693), An estimate of the degrees of the Mortality of mankind, drawn from curious Tables of the births and funerals at the City of Breslaw; with an attempt to ascertain the price of Annuities upon lives, *Philosophical Transactions…,* XVII, 596-610, 654-656.

JENNER, E. (1798). *An Inquiry into the Causes and Effects of the Variolae Vaccine, a Disease discovered in some of the Western Counties of England, particularly Gloucestershire, and know by the Name of the Cow Pox*. Sampson Low, London.

SÜSSMILCH, J. P. (1741), *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung desselben erwiesen,* Berlin. Existe una versión en francés de 1998, traducida y anotada por J. M. Rohrbasser, y publicada por l'Institut National d'Études Démographiques.

TREMBLEY, J. (1799). Recherches sur la mortalité de la petite vérole. *Mém. De l'Acad. Roy. Des Sciences pour 1796* pp. 17-38. Berlin.

Capítulo 5

Maurice Halbwachs y la estadística matemática

JOSÉ M. ARRIBAS MACHO
Departamento de Sociología I, UNED

Estadística y Ciencia Social

¿Porqué estudiar la obra de Maurice Halbwachs cuando es de todos conocido que las aportaciones fundamentales de la estadística matemática se producen en el campo de la escuela británica? Quizás por su proximidad con la Sociología, o mejor, por la destreza con la que introduce la estadística en los temas sociales. Bien cierto es que Maurice Halbwachs es mucho más conocido por sus aportaciones a la sociología que por sus trabajos como estadístico[1], pero sus análisis sobre las condiciones de vida de la clase obrera, sus trabajos demográficos, o su revisión de los estudios de Durkheim sobre el suicido, son suficientes para considerarle entre los pioneros de la estadística social aplicada. Su colaboración con Simiand y su familiaridad con la sociología empírica norteamericana (además de una situación biográfica privilegiada que le permite vivir el nacimiento de la estadística matemática), le convierten en un personaje clave para entender las relaciones entre la estadística y la sociología europeas. Probablemente, su muerte truncó una de las posibilidades de integración más interesantes de la sociología teórica continental y la sociología empírica anglosajona.

Halbwachs nace en 1877, y termina los estudios de filosofía en la Escuela Normal Superior de Paris en 1898. Entre 1905 y 1909 realiza estudios de economía, derecho y matemáticas en la Sorbona, periodo en el que conoce a Durkheim y Francois Simiad. En 1909

---

[1] Tal vez las obras más conocidas de Halbwachs son las relacionadas con la memoria. El trabajo de Olivier Martín es el único, que conozcamos, donde se aborda de forma independiente su contribución a la Estadística. MARTIN O. *Raison statistique et raison sociologique chez Maurice Halbwachs,* en Revue d' Histoire des Sciences Humaines (1999) pp 69-103.

lee una tesis de doctorado sobre las expropiaciones y el precio del suelo en Paris, y realiza una estancia en Berlín donde estudia economía política y marxismo. Su interés por la estadística se remonta a ese período, y es ya bien patente en los primeros estudios sobre la clase obrera, así como en su acercamiento al cálculo de probabilidades. Su tesis complementaria de doctorado, (1912, en la Facultad de Letras de la Universidad de Paris) versó sobre Quetelet y la estadística moral; en ella criticaba la teoría del hombre medio[2] y hacía una presentación del cálculo de probabilidades a partir de textos de Quetelet (Instructions populaires sur le calcul des probabilités) y Borel (Elements de la teories des probabilites).[3] Años más tarde, en 1924, y en pleno auge de la estadística matemática, publicará junto a un prestigioso matemático como Maurice Frechet[4], un tratado de divulgación sobre el cálculo de probabilidades (Le calcul des probabilités à la portée de tous).

Entre los estudios que han aparecido sobre Halbwachs durante los últimos años, tal vez sea en el de Olivier Martin (Raison statistique et raison sociologique chez Maurice Halbwachs, 1999) donde mejor se abordan sus puntos de vista sobre la estadística. Martin hace un excelente repaso de todos sus textos, aunque incide demasiado en las críticas sobre el uso desafortunado de la estadística, sin diferenciar los diferentes períodos en los que habla Halbwachs, y sin tener en cuenta que ciertamente, durante el período de gestación de la estadística matemática se producen abundantes errores y abusos metodológicos. Tal vez por ello, el monográfico de la Revue d'Histoire des Sciences Humaines pude haber contribuido a dar una imagen algo negativa de su interés por la estadística[5]. No obstante, es un trabajo fundamental para iniciarse en su pensamiento estadístico y familiarizarse con algunas de las ideas fuertes, por ejemplo que los métodos estadísticos sin teoría no son nada, que en ocasiones, el papel de la estadística debe ser modesto, los riesgos de simplificación que entraña el uso de estadísticos y la formalización, la importancia de los elementos subjetivos a la hora de definir hechos o grupos sociales etc.

De las encuestas de presupuestos familiares al índice del coste de la vida.

El interés de Halbwachs por la estadística aplicada comienza con la cuestión social, y se fundamenta en los estudios sobre condiciones de vida de la clase obrera. Las estadísticas sociales que ocuparon el interés de políticos y sociólogos durante toda la segunda parte del siglo XIX, continúan haciéndolo durante toda la primera parte del XX, y Halbwachs realiza en este campo dos minuciosos, aunque diferentes, estudios. Ambos pertenecen a épocas algo distanciadas y están lógicamente condicionados por la propia evolución de la disciplina estadística y el modo como se producen los datos estadísticos[6] .

---

[2] HALBWACHS, M. La théorie de l' homme moyen. Essai sur Quetelet et la statistiqu moral", Felix Alcan, Paris, 1912., p.14.

[3] También cita a Poincare: "Science et methode" y a Bertrand: Calcul des probabilités.

[4] Maurice Frechet mantuvo múltiples contactos con estadísticos y matemáticos españoles. Su primer viaje lo reliza en 1942, (véase BARBUT M.: Un épisode insolite des relations scientifiques franco-iberiques, III Congreso Internacional de Historia de la Estadística y de la Probabilidad, Madrid, 7 y 8 de julio de 2005) y fue invitado de honor en la Universidad de Madrid en 1950, poco antes de la creación de la Escuela de Estadística.

[5]En palabras de O.Martín, para Halbwachs, las matemáticas no sirven más que para establecer hechos que el sociólogo debe enseguida interpretar y explicar, y que deben ser emplazados en su contexto social preciso. Revue d' Histoire des Sciences Humaines (1999) pp 69-103.

[6] Es a partir de mediados del siglo XIX cuando comienza a generalizarse la existencia institutos centrales de estadística.

El primer trabajo, La classe ouvrière et les niveaux de vie, es la tesis principal de la Facultad de Letras -que lee en 1912, con 35 años-, y en la que trata la unidad de la clase obrera desde los hábitos de consumo. Halbwachs se propone estudiar las necesidades y los gastos de esa clase social, pero a diferencia de los autores que se han centrado en la búsqueda de regularidades y leyes estadísticas, nuestro sociólogo analiza la variabilidad de los grupos y la diversidad de las condiciones sociales relacionándolas con el consumo[7]. Para ello utiliza dos encuestas alemanas: una confeccionada por la oficina imperial de estadística, y otra elaborada por la Unión de trabajadores del metal de Alemania (Haushaltungsrechnungen von Metallarbeitern). Allí se plantea el fundamento de la jerarquización social en las sociedades industriales partiendo del consumo como práctica social, y no de las condiciones del trabajo, como se había hecho hasta entonces. Dedicará el libro primero a justificar la pertinencia del estudio de la clase obrera, mientras que en el segundo se centra en el consumo desde el punto de vista empírico (los gastos). En el tercero reflexiona sobre las tendencias del consumo y establece algunos elementos para una "teoría sociológica de las necesidades"[8].

Halbwachs también incorpora un amplio capítulo sobre metodología (preparación de los cuadernillos, duración de la encuesta, el número de presupuestos, los diferentes tipos de gastos, medias ficticias y medias válidas, la expresión del tamaño de las familias..) que al parecer constituían un requisito básico de las tesis realizadas en el entorno durkheimniano[9]. En él se plantea el examen crítico de los datos, la posibilidad de que esas anotaciones estadísticas "no se correspondan con la realidad", por ejemplo, que el obrero que rellenó el cuadernillo disimuló algunas compras y añadió otras por razones de prestigio social, morales, etc., y ello, a pesar de que las encuestas le ofrecen garantías porque estaban dirigidas por importantes organismos oficiales o colectivos. La administración alemana contaba ya con suficiente prestigio y autoridad para garantizar el buen uso de los cuestionarios, no en vano, la invitación a participar en la recogida de datos se hacía al servicio de "la administración y la ciencia" -de forma casi gratuita-.[10] En las conclusiones, Halbwachs, se reafirmaba en las viejas hipótesis durkheimnianas: de todo este análisis queda claro que la naturaleza de las necesidades esenciales, así como su cantidad, se explican por la sociedad.

Hay, no obstante, una gran diferencia entre este trabajo y el que publica en 1933 (L'évolution des besoins dans les classes ouvrières), el más interesante, desde nuestro punto de vista, puesto que refleja los cambios que se han producido en el tratamiento empírico de los fenómenos económicos. En menos de una centuria (1870-1950), el análisis económico se ha transformado radicalmente, incorporando una serie de herramientas matemáticas y estadísticas que aún no estaban en uso durante la elaboración del primer estudio. En este segundo trabajo -publicado veintiún años más tarde-, hay un verdadero salto cualitativo; para empezar, Halbwachs ya no se limita a analizar los resultados de dos encuestas de un mismo país, sino que dispone de suficiente información para realizar un estudio comparativo internacional, pues la información que se ha ido produciendo en las oficinas centrales de los

---

[7] Según los oficios, las ramas de la industria, las regiones, etc. En el estudio de los años treinta, tampoco pierde de vista "la unidad de la clase obrera" aunque ahora las comparaciones son de tipo internacional.
[8] El paso de una teoría objetiva del valor trabajo a una teoría subjetiva representada por las corrientes de la utilidad marginal, desplazará el trabajo de las preocupaciones de los economistas para centrarse en el consumo. Keynes y otros tantos economístas de la época entrarán en estos debates. Véase ALONSO L.E. La producción social de la necesidad, en Estudios Sobre el Consumo nº
[9] Véase TOPALOV "Experiences sociologiques" les faits et les preuves dans les thèses de Maurice Halbwachs (1909-1913), en Revue d'Histoire des Sciences Humaines, nº 1, 1999.
[10] Participaron en la recogida de ese material 218 empleados de clase media y maestros, además de 522 obreros o empleados.

países industrializados es muy abundante. A partir de la primera guerra mundial, los estudios sobre presupuestos familiares se multiplicaron en todos estos países, en España, por ejemplo, los trabajos que habían comenzado con la Comisión de Reformas Sociales recibieron un gran impulso durante los años veinte con el Instituto de Reformas Sociales, el organismo que se encargará de poner en marcha los estudios sobre el "coste de la vida del obrero" [11]. El IRS español comenzó muy pronto a elaborar series temporales e informaciones estadísticas que condujeron, aunque después de la guerra, al índice del coste de la vida. Del mismo modo, la etapa más espectacular de innovación estadística comenzó en Alemania a partir de 1924, de la mano del profesor Erns Wagemann, en la Oficina de Alimentación[12]. Allí confluyeron las estadísticas oficiales con una nueva forma de economía conocida como Konjunkturforschung, y que hoy podríamos llamar análisis de coyuntura de ciclos económicos, que está íntimamente ligada a la aparición de la econometría y la estadística matemática.[13] Instrumentos como balanza de pagos, tasa de desempleo, índice de precios, producto nacional, etc., que comienzan a diseñarse durante esos años, se convirtieron muy pronto en las piezas fundamentales de una economía que necesita la producción continua de datos estadísticos.

En el estudio de La evolución de las necesidades en las clases obreras, Halbwachs maneja la información que aparece en los anuarios estadísticos de países como Estados Unidos, Canadá, Francia, Argentina, o las encuestas de la Oficina Internacional del Trabajo realizadas en diferentes ciudades del mundo, por ejemplo Barcelona[14]. Este segundo estudio está ya marcado por un lenguaje en el que se habla de poder de compra, evolución de precios, coste de la vida, y en concreto del índice del coste de la vida, un lenguaje moderno y muy similar al

[11] Véase ARRIBAS JM Y VALLEJOS A (20..).Para una introducción historiográfica general al IRS y a la CRS y una visión de conjunto de su actividad pueden verse los trabajos de Palacio Morena (1988) y de Calle Velasco (1989). En España, la producción de estadísticas sociales se realiza en el seno de la Comisión de Reformas Sociales y del Instituto de Reformas Sociales (1903-1924), Instituto que acabaría asumiendo como fin propio orientar técnicamente la intervención del Estado en los asuntos que afectan a las masas obreras, en las relaciones entre capital y trabajo. El Instituto (IRS) supone la consolidación de la empresa iniciada en 1883 con la Comisión que acabó llamándose 'de Reformas Sociales' (CRS) y que en principio fue una extraña comisión ad hoc encargada de "practicar una información sobre el estado y necesidades de la clase obrera".. El Instituto de Reformas Sociales empieza a regularizar informaciones y a establecer series estadísticas que permiten conocer la evolución temporal de ciertos fenómenos anómalos, como las huelgas y los accidentes laborales, o agregados, como el coste de la vida del obrero.

[12] Véase TOOZE, (2001). La economía aparecerá por primera vez, contemplada como un sistema independiente y separado de "lo social", incluso de lo "político", convirtiéndose en algo mensurable y, por tanto, en objeto científico. En Alemania, esta nueva teoría macroeconómica fue el contexto de un amplio e innovador programa de investigación estadística fuertemente esponsorizado por la república de Weimar, lo que sitúa a Alemania, según Tooze, entre los países que constituyen el crisol de la moderna estadística económica (Estados Unidos, la Unión Soviética, Gran Bretaña, Suecia y Holanda).

[13] Véase ARRIBAS J.M. Les débuts de la statistique mathématique enn Espagne (1914-1936) Mathématiques et Sciences Humaines, nº 166, CAMS, EHESS, Paris.

[14] Las encuestas estaban patrocinadas por la Fundación Ford para comprobar los estilos de vida de los obreros de Detroit con los de otros países europeos. La elección de Barcelona, esta relacionada con el hecho de que Ford comienza a ensamblar coches en esa ciudad. En España, la inspección de trabajo produce abundante información estandarizada. Las informaciones de los inspectores de trabajo se fijan en las infracciones de las leyes que regulan el trabajo de mujeres y niños y el trabajo nocturno, pero también pretenden la elaboración de un 'censo de la población obrera en España'. Para ello registran datos de los centros de trabajo, clase de industria, localización territorial, número de obreros empleados, distinguiendo por sexo y edad, salarios máximos, mínimos y medios. Los anuarios que bajo el nombre de 'Estadísticas de huelgas' se publican desde 1904 hasta 1922 combinan un exhaustivo conjunto de tablas con unos informes narrativos más o menos extensos en donde se da cuenta a veces de manera pormenorizada del desarrollo y de la resolución de las huelgas más notables. Véase ARRIBAS Y VALLEJOS (2002).

que se emplea en la actualidad y que Halbwachs maneja con soltura. En relación con los trabajos de Wagemann sobre la previsión de las crisis dice lo siguiente:

> *"A diferencia de los estadísticos de Harvard, el Instituto alemán no se detiene a observar de forma regular y a representar por sistemas de curvas las variaciones del curso de los hechos, de los precios, de la tasa de descuento, etc. Sino que se esfuerza en penetrar con más profundidad en la estructura de la organización industrial, de medir los movimientos de la producción, los stocks, las salidas, el consumo, es decir, de utilizar en función de las previsiones, un número mayor de "barómetros económicos*"[15] .

Resalta el carácter "intensivo" de la nueva encuesta alemana, aunque se lamente de que solo facilita información de un año, lo que hace difícil las comparaciones con el estudio realizado veinte años antes. Si utiliza las encuestas realizadas en los EEUU durante el período 1895-1930, es porque le permiten considerar la evolución del período:

> *"Es cierto que el método aplicado en América para el establecimiento de los presupuestos familiares no es el mismo que en Alemania, y parece menos exacto. No se apoya sobre las anotaciones de gastos realizadas día a día por las familias, sino que los encuestadores les plantean preguntas y escriben las respuestas que tratan sobre períodos más largos*."[16]

En realidad, lo que está sucediendo es que los americanos, bajo la influencia de sus colegas británicos, están ya ensayando la encuesta representativa que va a generalizarse después de la II Guerra Mundial en todas las oficinas centrales de estadística del mundo.

Desde el lado metodológico, a Halbwachs le interesa la construcción de series temporales, la comparación de la evolución del gasto realizado por las familias obreras con la situación de los ciclos económicos. Aunque la información sobre presupuestos familiares es todavía parcial y fragmentaria, los datos de producción y comercio –bastante más completos- le permiten hacer estimaciones sobre consumo por habitante (consommation par tête): « *Es posible evaluar el gasto correspondiente para cada año, multiplicando las cantidades por los precios, así como reconstruir y seguir en su evolución, una parte importante del presupuesto de gastos de un hogar medio*"[17]. Para el caso de otros consumos, pan, carne, vino, etc., construye tablas de números índices, y visualiza sobre la base 100 del primer año, los incrementos relativos de los períodos sucesivos. De este modo, aparece una imagen didáctica de la evolución de la serie, la tendencia, al alza o a la baja, los picos, etc., tablas que, por supuesto, acompaña de su correspondiente gráfica. Otra pesquisa metodológica fundamental es la estructura del gasto y su evolución, es decir, el peso relativo de cada uno de los componentes (alimentación, vestido, alojamiento, ocio, etc.) en el gasto total, así como la aparición de nuevas necesidades en la clase obrera (le interesan los mecanismos de producción y su evolución) y el consumo de los nuevos objetos fabricados en serie que conoce bien por las encuestas americanas (El automóvil es, tal vez, el nuevo objeto de consumo que más le fascina).

---

[15] Cita el texto de WAGEMANN en francés, publicado por Alcan en 1932: Introduction à la theorie du mouvement des affaires. Halbwachs, 1933, p.VI. Para los barómetros económicos, véase ARMATTE, M.(1992) Conjonctions, conjocture et conjecture. Les barometres économiques (1885-1930), Histoire et Mesure.

[16] HALBWACHS, M. (1933) L'évolution des besoins dans les classes ouvrières, p.VII.

[17]Op cit, p.128.

Qué representa el "índice del coste de la vida", se pregunta[18]. Un estadístico que calcula todos los meses la Statistique general de la France -como lo hace ya la mayor parte de los países industriales-, es una lista de objetos de consumo bien definidos en cantidad y calidad. Una definición que M. Huber, presenta en la XVII sesión del Instituto Internacional de Estadística (1927) le sirve de punto de partida. En ella se afirma que se trata de "seguir las variaciones del poder de compra de la moneda en lo que concierne al conjunto de bienes y servicios necesarios para asegurar un cierto género de vida"[19] El boletín de la Statistitique Générale de la France et du service d'observation des prix precisa un poco más y recomienda

> *"tomar por base el cálculo del presupuesto ordinario de una familia obrera de 4 personas y de aplicar a las cantidades fijas de diversos objetos de consumo que entran en ese presupuesto, los precios unitarios variables observados en épocas sucesivas, de manera que el índice represente el gasto de una familia manteniendo constante su género de vida"*[20].

Halbwachs observa el modo como en los diferentes países industriales se calcula el peso de los gastos (hoy diríamos de los bienes y servicios consumidos) que intervienen en el índice. Toma en consideración Canadá, Francia, Inglaterra, Alemania, Italia, Estados Unidos y un organismo internacional tan significativo como el Bureau Internacional du Travail. Pero nos interesa detenernos en el sistema anglosajón porque es de donde va a surgir la encuesta representativa[21]

Inglaterra facilita los pesos (coeficientes), precisa Halbwachs, utilizados para calcular el índice de la alimentación correspondiente al gasto medio de 1.944 familia obreras, y cuyos presupuestos familiares fueron recogidos en 1904 por el Board of Trade. Halbwachs considera esos resultados cuestionables por el hecho de que solo dieran información de una semana, pero reconoce que para establecer un presupuesto-tipo, estas acciones pueden ser suficientes. Los norteamericanos en cambio, consideraba dos índices diferentes, el del Bureau of Labor Statistics, y el del Nacional Industrial Cenference Board, en tanto que las proporciones atribuidas a cada gasto se determinaron de acuerdo a una encuesta realizada en 1917-19 para el primero, y según una encuesta realizada en 1901-1903 para el segundo. Añade un malicioso comentario, en el sentido de que en un país donde todo cambia tan rápidamente, no es que los estadísticos americanos hayan tenido el mal gusto de inclinarse por la inmovilidad, estandarizando las necesidades por tiempo indefinido, sino que tratan de comprobar como varía un tipo de vida cuando se le supone constante.

---

[18] Para la historia de los índices de precios véase ARMATTE M. Indice des prix: histoire et controverses. En SANTOS DEL CERRO J. y SECADES M. (coord.). Historia de la probabilidad y la estadística II, A.H.E.P.E , Delta Universidad, 2004.
[19] Halbwachs (1933), p.1
[20] Ibidem.
[21] En Canadá calculaban tres índices del coste de la vida y no se apoyaban sobre encuestas de presupuestos familiares, sino en la estimación del consumo medio por habitante. En vísperas de la I Guerra Mundial, una comisión del Ministerio de Trabajo francés intentó organizar una encuesta sobre ingresos y gastos de obreros y empleados, al estilo americano, con cuestionarios. En 1922 se creó una Comisión central que ponen en marcha un sistema de encuestas por cuestionario, aunque los presupuestos tipo resultantes fueron extremadamente diferentes y la Comisión optó por uniformizar el peso de los componentes del índice de acuerdo con una estructura puramente teórica, que asignaba a la alimentación un peso del 60%; al vestido, un 15%; a la calefacción y limpieza, un 5%; y al alojamiento, un 12%. El 8% restante fue asignado a gastos diversos.

“Las causas del suicidio”

Las causas del suicidio fue un trabajo concebido como la actualización del famoso texto de Durkheim “El Suicidio”, la obra metodológica que está íntimamente ligada a “Las Reglas del método sociológico”. En la disputa con otras corrientes psicológicas y sociológicas para establecer el “método” de la ciencia social, Durkheim mostraba a partir de datos estadísticos sobre suicidios, que un hecho, en apariencia individual como quitarse la vida, es un hecho social, y que ello quedaba demostrado por las regularidades de las tasas de suicidios de determinados países, épocas y grupos sociales. La cifra social, expresaba un cierto estado del “alma colectiva”, un cierto estado de la opinión, y las regularidades estadísticas no eran sino la expresión de esas fuerzas sociales colectivas que en la teoría durkheimniana empujan al individuo a obrar.

El éxito del trabajo de Durkheim sobre el suicidio hizo posible que la sociología contara por primera vez con un estudio canónico de sociología empírica donde se hacían explícitas las reglas del método científico. Allí quedaba patente que los actos individuales respondían a causas sociales y se establecía la gran divisa metodológica durkheimniana:“los datos estadísticos expresan la tendencia al suicidio de cada sociedad”. Con gran maestría, Durkheim manejaba el método de la experimentación indirecta fundada sobre las variaciones de las tasas de suicidio al ponerlas en relación con otros parámetros de la vida social como la religión, el estado civil, la profesión etc., pero con el paso del tiempo, la información estadística se fue quedando obsoleta, y en el trabajo quedaron patentes algunas lagunas y problemas metodológicos. Según se desprende de la introducción a “Las causas del suicidio”, es el propio Marcel Mauss quien solicita a Halbwachs una revisión y actualización del texto para comprobar si las nuevas informaciones estadísticas confirman, o no, las tesis expuestas por el maestro.

Las razones eran de dos tipos: en primer lugar, las fuentes estadísticas que había recopilado el propio Mauss para “El Suicidio” respondían a un estado de la estadística anterior a 1891, y además, reconocía que los métodos estadísticos se habían desarrollado bastante desde que el maestro publicara aquel trabajo. A fin de cuentas, Durkheim no poseía formación matemática -su formación era clásica-; y su famosa tipología de las causas del suicidio se apoyaba, en gran medida, en sus conocimientos etnográficos. En la introducción, Mauss afirma que la idea primera fue actualizar el trabajo de Durkheim con un capítulo suplementario que aportara los nuevos datos estadísticos, pero añade que Halbwachs, “poco a poco se ha ido sintiendo obligado a realizar nuevas investigaciones, a plantear nuevos problemas y a presentar los hechos desde un nuevo aspecto” ¿Qué aspecto nuevo era ese? No la teoría sociológica general, pues el suicidio continúa siendo para Halbwachs un hecho social y Halbwachs es muy respetuoso con el maestro, pero si incorpora novedades importantes como veremos a continuación.

Para empezar, Halbwachs da mucho más valor al especto psicológico individual, en realidad considera el suicidio un hecho individual y social, a la vez; establece clasificaciones de suicidios por los motivos, recuperando una vieja categoría hegeliana, y no tanto por las causas. Pero además, incorpora nuevas herramientas de análisis estadístico, tal vez, la diferencia fundamental de ambos estudios provenga, como en los estudios sobre el consumo de la clase obrera, de la época en la que fueron hechos. La de Durkheim, es una época que esta a la búsqueda de las leyes causales que han de dar cuenta de los fenómenos sociales, al modo como lo hacen las leyes de la física clásica, y su referente es Quetelet. La de Halbwachs, por el contrario, es una época que vive el cambio de paradigma que representa el nacimiento de la física cuántica, junto a la recuperación de la probabilidad y el azar como fundamentos científicos[22].

Christian Baudelot y Roger Establet, en un reciente e interesante trabajo[23], destacan como Halbwachs abandona la famosa tipología de suicidios de Durkheim (egoísta, altruista y anómico), y el concepto de anomia tan querido por el maestro, es sometido a una revisión tan profunda que los cambios de las sociedades industriales son considerados -a diferencia del maestro- como mayor grado de complejidad y no como "desorden". Halbwachs Llega a decir que la vida social moderna es más normativa que la antigua debido a que está dominada por las leyes del mercado[24].

Otra de sus aportaciones, como ya hemos señalado, es la incorporación de los objetivos personales del sujeto, puesto para Halbwachs, la sociedad existe en el conjunto y en los individuos. La ruptura de la frontera metodológica individuo/sociedad es algo que comparte con la visión de Mauss sobre el "hecho social total" (la tesis inacabada de Mauss versaba sobre "le prier", el rezo, una acción individual y social, a la vez); porque la sociedad está en cada uno de sus individuos, del mismo modo que las propiedades de un elemento químico, por ejemplo el uranio, se encuentran en cada una de sus partes. La sociedad no existe fuera de los individuos que la encarnan: los sentimientos de familia, las prácticas religiosas, la actividad económica toman cuerpo en las creencias y en las costumbres que unen las existencias individuales[25].

Además de contar con mayor y mejor información estadística[26], Halbwachs reconoce como hacía Mauss que los métodos estadísticos han progresado: "Ya no se limitan los estadísticos a calcular las medias, las proporciones o porcentajes. Un sociólogo americano M.John Rice Miner se sorprendía recientemente de que no se utilice todavía en los estudios de suicidios los procedimientos estadísticos modernos, cálculo de las desviaciones, índices de correlación, de dispersión, etc.". Debido a ello, -aunque le parece algo artificial y arbitrario- introduce el concepto de desviación, y sobre todo, una nueva herramienta como el coeficiente de dispersión ponderado[27] de las tasas de suicidios, que utiliza para los datos que van desde 1836-45 hasta 1922-25. Información que le sirven para contradecir la tesis de un especialista como Morselli que había asegurado que en los Estados "civilizados" el suicidio aumenta desde el comienzo del siglo XIX con una rapidez superior a la de la población y la mortalidad.

---

[22] Los años veinte son una época de extraordinaria vitalidad intelectual que algunos han llegado a considerar la más importante desde la época del pensamiento griego; sirvan como ejemplo las interpretaciones elaboradas por el Instituto de física de Copenhague entre 1925-27 en torno a Bohr sobre el espacio la causalidad y el tiempo, o los artículos que pueden encontrarse en un mismo número de la Revue Philosofique francesa sobre la noción de causa (Meyerson), sobre física cuántica (Goblot) o sobre estadística (Halbwachs).

[23] BAUDELOT C. Y ESTABLET R. Suicide:changemente de régime. En JAISSON M. Y BAUDELOT, C. Maurice Halbwachs, sociologue retrouvé.Editions Rue d'Ulm, 2007.

[24] Ibidem, p. 25

[25] BAUDELOT (2007) p.27.

[26] Los datos de Durkheim estaban iban desde 1840 hasta 1890-91 y Halbwachs los califica de valor desigual. Un Prusia la estadística del suicidio no está completa sino a partir de 1883, en Inglaterra comienza en 1856, y en Italia en 1864. Para el conjunto del imperio alemán solo se disponen de cifras a partir de 1881.

[27] Halbwachs explica lo que es la dispersión: Dada una serie de números, se puede calcular la media. Pero una misma media puede representar tanto a una serie de números próximos como a una serie de números alejados. La dispersión representa el alejamiento, más o menos grande de los términos de una serie, unos en relación a otros, y en relación a su media. He aquí como se la puede medir: se calcula en principio, la desviación positiva o negativa, de diversos elementos de la serie en relación a su media. Se hace la suma aritmética de esas desviaciones (es decir, sin tener en cuenta los signos + y -) Se multiplica, enseguida esta suma por cien, y se la divide por la suma de los números de la serie. En nota a pie de página, añade que cuando las desviaciones son muy pequeñas, se calcula la desviación cuadrática media, también llamada Standard deviation. (Halbwachs, 1930, 83. edición electrónica Mme. Marcelle B ergeron)

Con sus coeficientes de dispersión, Halbwachs demuestra que en una primera fase (1845-1870) el suicidio aumenta y que en una segunda, que se prolonga hasta 1913, disminuye.

Contrariamente a los estadísticos anteriores, nuestro sociólogo fija su atención en el número de suicidios y su evolución al modo como lo hacen las Administraciones Públicas, lo que le permite rectificar algunas de las conclusiones del maestro: las observaciones realizadas sobre una duración más amplia nos han permitido rectificar notablemente las conclusiones y previsiones de Durkheim; perspectiva que se inscribe dentro del tránsito de la política del s. XIX[28] a la política del s.XX, en el contexto de una política sustentada por una actividad científica que produce soluciones alternativas ante problemas aislados; en otras palabras, los trabajos de Halbwachs se inscriben en una forma de abordar las cuestiones sociales mucho más práctica (que requiere la intervención del Estado en asuntos concretos) y bastante más técnica (requiere la intervención de especialistas y científicos del Estado), que va a tener en la Administración británica a uno de sus principales exponentes.

La introducción de la probabilidad y algunas cuestiones metodológicas.

La introducción de la probabilidad en las ciencias sociales alcanza la madurez cuando se desarrolla la técnica y el método de la encuesta representativa. Sigue, no obstante, caminos diferentes, por ejemplo Halbwachs en su artículo de 1923 utiliza la vía de validar la distribución de datos estadísticos censales, porque todavía no ha desparecido del todo la búsqueda de leyes causales. Según ese planteamiento, un determinado fenómeno puede distribuirse al azar o, por el contrario, distribuirse de una determinada forma que justifique la existencia de una ley (por ej., las leyes de la herencia de Mendel).

En el caso de la aplicación de la probabilidad a la teoría muestral, aparte de los cálculos realizados por Laplace sobre la población de Francia y las aplicaciones de la curva de errores en astronomía, los estadísticos esperaron hasta finales del siglo XIX para celebrar reuniones científicas destinadas a debatir sobre la utilización del método representativo, con la probabilidad como base de legitimación. A partir de 1895, el estadístico noruego Kiaer comenzó a presentar comunicaciones en las reuniones del Instituto Internacional de Estadística (Berna, 1895; San Petesburgo, 1897; Budapest, 1901)[29] en las que se discutió acerca de la validez de los datos estadísticos obtenidos con pequeñas muestras y procedimientos aleatorios, y en 1909, Francis Isidro Edgeworth presentó por primera en un coloquio del Instituto Internacional de Estadística (París), un amplio trabajo matemático sobre la aplicación del cálculo de probabilidades a la encuesta estadística[30]. Aunque por esas fechas, el método estaba ya muy avanzado, no será hasta el coloquio celebrado en Roma, en 1925, cuando los estadísticos aceptaron oficialmente la plena validez del método aleatorio probabilístico.

En Inglaterra fueron las encuestas sobre las condiciones de vida de la clase obrera las que propiciaron las investigaciones sobre este método[31]. En mayo de 1887, Charles Booth había

[28] Véase ARRIBAS J.M. El desarrollo de la ciencia estadística durante el siglo XIX en INE, "150 aniversario de la Estadística oficial española", 2007.

[29] Es en esta reunion cuando Kiaer muestra la carta de adhesión a su método del USA Department of Labour

[30] EDGEWORTH F Y On the application of the calculus of probabilities to statistics IIS, XII Session, Paris, 1909, pp505-551.

[31] Habría que señalar también los importantes avances de la estadística rusa durante el período 1885-1917 en los zemstos. La reflexión de los estadísticos rusos y la calidad de las encuestas durante ese período por la asignación óptima utilizando estratos que A.G. Kovalevskij presenta en 1924, diez años antes que Neyman. Véase al

leído en la Royal Statistical Society un informe sobre las condiciones sociales y económicas de la pobreza, que había sido redactado a partir de una encuesta realizada en el East London, y cuyos resultados fueron publicados entre 1889 y 1903 (en 17 volúmenes). La encuesta de Booth utilizaba un método indirecto: un cuestionario que los inspectores educativos se encargaban de rellenar al estilo de las encuestas desarrolladas en España por la Comisión de Reformas Sociales, mientras que las encuestas de John Hilton y Arthur Bowley realizadas entre 1914 y 1924[32], se hicieron mediante entrevistas directas, casa por casa, y lo que es más destacable, la selección de los entrevistados se realizó mediante procedimientos aleatorios.

Halbwachs publicara en 1923 un artículo en el número 96 de la Revue philosophique en el que trata el papel de la probabilidad en la investigación social, pero no aborda directamente la encuesta estadística (en el mismo número, Meyerson habla sobre la noción de causa y Goblot sobre física cuántica). El título de Halbwachs es Experimentación estadística y probabilidad[33] y comienza preguntándose si existe alguna diferencia entre el estadístico que analiza columnas de cifras y gráficos, y el físico que organiza un experimento. A primera vista, el estadístico se sitúa entre aquellos que trabajan con libros y anuarios, sin que parezca prudente atribuirle poder de decisión sobre los hechos que estudia; se le podría comparar con el historiador, dice Halbwachs, porque este no ejerce ninguna influencia sobre los hechos del pasado. Elige como ejemplo social, la influencia que tiene la edad sobre la mortalidad en un grupo: si el organismo fuese un agregado químico relativamente simple, sería sencillo descomponerlo en sus elementos y analizar las relaciones:

> "*podríamos proceder como el químico que somete un compuesto conocido a reacciones que le descomponen. Pero cuales son los elementos de un organismo, cuales son sus relaciones, qué acciones se ejercen sobre él en cada edad? El problema es infinitamente más complejo, no solo porque se trata de un organismo sino porque además, todos los miembros de una especie de la misma edad no son idénticos*"[34].

El investigador social procede de un modo algo diferente:

> *"Uno debe limitarse a observar lo que pasa, es decir, sin ejercer una acción sobre el grupo, debe limitarse a contar la proporción de hombres de una misma edad que mueren cada año. Si esta proporción es constante, llegamos a un resultado, pero por otro método completamente distinto, aparentemente, al del experimentador*"[35]

No obstante, ¿se trata de un método científico diferente? se pregunta Halbwachs. Su opinión, como la de Simiand es negativa y recurre a la tasa de paro mensual en una serie de años para ilustrarlo[36]. La comparación es compleja porque se mezcla la variación de un período anual, según el mes o las estaciones, con otra de período más largo que mide la tendencia a través de diferentes años. Con procedimientos estadísticos apropiados, puede eliminarse la variación interanual, de manera que se aísle la variación interior al año. De la

---

respecto los interesantes trabajos de MARTINE MESPOULET Du tout à la partie. L' âge d' or du sondage en Russie (1885-1924). Revue d' etudes comparatives Est-Ouest, vol 31, nº 2, pp5-49.

[32] Véase ARRIBAS J. M: Presentación del texto de Arthur L. Bowley La aplicación del muestreo a los problemas económicos y sociológicos. Empiria, nº 5, 2002, pp195-204-

[33] HALBWACHS M. (1923) L'expérimentation statistique et les probabilités . Revue Philosophique de la France et de l'Étranger. Juillet-décembre 1923. p340-371.

[34] Op. Cit. P.340.

[35] Ibidem

[36] Se trata de un ejemplo de Simiand, el verdadero maestro de Halbwachs en cuestiones de estadística. El ejemplo lo extrae de: SIMIAND F. Statistique et experérience. Remarques de méthode., Riviere, 1922.

misma manera puede eliminarse la variación estacional y tener solo en cuenta las variaciones interanuales, de modo que pueda estudiarse la relación que cada una de estas variaciones con cualquier factor. ¿En qué procedemos de manera diferente al físico, se pregunta Halbwachs, cuando se descompone un fenómeno en sus elementos para examinar su efecto sobre el resto? Desde luego, no en lo esencial, otra cuestión es la diferencia que exista entre lo grupal y lo individual en la sociedad, aunque todo depende de la perspectiva que queramos adoptar:

> *"Para el grupo se obtiene una cifra, o cifras, que miden exactamente la frecuencia relativa de parados en tal o cual condición con el mismo resultado positivo que en cualquier otra ciencia. Para un miembro del grupo no procedemos de igual manera: poniendo en relación el número de parados con el número de obreros del grupo se puede calcular lo que se llamará la probabilidad (chances) de estar parado de un obrero cualquiera. Pero esto no es un resultado positivo, en el sentido de que no se si tal obrero terminará en el paro o no. Es por esta razón que algunos lógicos han creído necesario distinguir entre probabilidad e inducción, de modo que solo en ésta última se pueden fundamentar las leyes, mientras que sobre la probabilidad no se pueden fundamentar más que previsiones más o menos creíbles."*[37]

Y aquí se apoya en J.Venn, en concreto en su Logic of chance de 1888. El método inductivo al que se refiere Halbwachs, nos remite a la encuesta estadística representativa aunque en el artículo no trate directamente el asunto.

Pero veamos como se desarrolla su razonamiento: en física las leyes permiten hacer previsiones seguras, dice nuestro sociólogo, mientras que el uso de la probabilidad ha sido introducido -por ejemplo en los seguros- por razones de comodidad, de manera que "un grupo ficticio de individuos, supuestos idénticos, sustituye a un grupo natural dado de hombres de tal edad." Hay grupos sobre los que se puede establecer con exactitud que tal fenómeno se producirá, aunque no se pueda determinar en qué individuos concretos sucederá, del mismo modo que en física tampoco se conoce como se distribuye la temperatura en las partes más elementales de un cuerpo, o por donde romperá un cable cuando se le somete a una tensión determinada. Pero esa incertidumbre en el pequeño detalle no disminuye el valor de las leyes físicas, y su determinación tiene tanto interés en la física como en las cuestiones sociales. Los experimentos de la física que consisten en eliminar la acción de un factor (por ej. la temperatura) para estudiar otro (por ej., el volumen de un gas), se realizan en los procesos sociales mediante procedimientos estadísticos; por ejemplo, en el estudio del paro, para eliminar las variaciones estacionales, se calcula la media de paro de cada año.

Del mismo modo que el físico necesita conocer el manejo de los instrumentos que ha diseñado el ingeniero para sus observaciones, el estadístico necesita conocer el modo como los organismos de la administración del Estado organizan la información, pues con harta frecuencia aparecen lagunas, confusiones y clasificaciones que son fruto del trabajo de otros. Pone el ejemplo siguiente:

> *"Si los datos de paro no fueron recogidos mensualmente, por tipo de industria, región, o no se distingue a los parados de los vagabundos, de los minusválidos, etc. no podrá realizarse, ninguna acción sobre una masa que es un conjunto de hechos movedizos y confusos. El trabajo que corresponde al cálculo de las medias, consiste en recomponer sobre otro plano, sobre varios planos diferentes, el conjunto del que las estadísticas solo*

[37] HALBWACHS, M. L' experimentation statistiue et les probabilités, Revue philosophique, 96, 1923, pp.340-371.Ediciónelectrónica editada por Jean-Marie Tremblay, http://www.uqac.uquebec.ca/zone30/Clasiques des_sciences_sociales/index.html, p.4

nos representan las partes. Operación intelectual, y no física, como dice M. Simiand, porque nosotros reflexionamos sobre relaciones, es decir, sobre abstracciones"[38].

La conclusión del trabajo es que las operaciones estadísticas presentan los caracteres del método experimental y que están estrechamente ligadas a la teoría y al cálculo de probabilidades. A pesar de que las leyes estadísticas son diferentes de las del azar, y de que propone renunciar al empleo de las expresiones que designan resultados sin relación alguna con el "estudio experimental", su apuesta por la probabilidad no deja lugar a equívocos:

"El estadístico está obligado a cada instante a servirse del cálculo de probabilidades para analizar los objetos colectivos que observa y determinar aisladamente las variaciones de cada uno de sus elementos en sus relaciones con otros objetos colectivos o con otros objetos cualesquiera. En este sentido, el cálculo de probabilidades juega en estadística, aproximadamente, el mismo papel que "los instrumentos" en la experimentación físico-química"[39].

En 1933, escribe un artículo titulado "La ley en Sociología" en el que aborda otro de los nudos gordianos de la estadística y que va a ser objeto de sus críticas: la relación entre las leyes científicas y lo real; tema objeto de debate entre algunos investigadores que perfilan una nueva atmósfera intelectual en la que la comunidad estadística pasa a estar más interesada en la medición de los fenómenos sociales (por ejemplo el desempleo) que en su predicción.

Su punto de partida de lo que es una ley en ciencias sociales es Augusto Comte, pero su opción personal es Simiand y consiste en producir modelos empíricos:

Una ley parece ser una relación obtenida de una observación material, y, preferentemente cuantitativa, que se presenta bajo la forma de una proposición general (aunque sobre este punto se puede discutir) que es específica, es decir, que se establece entre términos homogéneos, del mismo orden o del mismo dominio: explicación de lo mecánico por la mecánica, de lo biológico por la biología, etc. (p.4).

En esto podríamos decir que Halbwachs está más cerca de Karl Pearson que de Edgeworth o Bowley, aunque su crítica hacia los economistas que denomina "escuela abstracta del valor", y que hace coincidir con los pioneros de la nueva economía matemática[40] es muy fuerte. A pesar de todo, seguimos creyendo que sus puntos de vista teóricos son muy avanzados para la época y que en lo esencial, son coincidentes con los utilizados por los teóricos de la utilidad marginal: partiendo de la idea de Durkheim de que los hechos económicos son sobre todo, creencias y hechos de opinión y se apunta a una teoría subjetiva del valor (lo que define los diferentes valores de los objetos, son las creencias y la evolución de esos valores es el resultado de la evolución de esas creencias) Sus críticas, por el contrario, están reservadas para los excesos en la aplicación de los nuevos métodos matemáticos.

[38] Op cit, p.8.
[39] Op. Cit, p.25
[40] Su posición es muy crítica con esta escuela: "Yo he leído a Cournot, he leído a Walras, he leído a Pareto, y debo decir que no me han enseñado gran cosa sobre la realidad misma, sobre los hechos mismos, yo no he tenido la impresión de que esto pueda ser incorporado a la ciencia positiva propiamiente dicha, y real" HALBWACHS, M. (1937) Le point de vue sociologique" Edición electrónica de Jean-Marie Tremblay. "Les clasiques des sciences sociales", pp15.

Dos años más tarde, el Centro Internacional de Synthèse,[41] organiza un seminario internacional sobre las aplicaciones de la estadística y en su exposición (La Statistique en Sociologie) y allí pone de manifiesto alguna de sus ideas más interesantes, por ejemplo que cualquier recuento no es una estadística[42], pues además es necesario que los datos presenten cierta consistencia. A diferencia de las ciencias físicas, donde si se suprimen algunos aspectos fundamentales durante el experimento, el fenómeno no se produce, en la estadística, las cifras se dejan siempre combinar con cifras[43]. Respecto a la utilización de curvas y gráficos, su posición es también crítica, pues vuelve a plantearse el problema de la relación entre los modelos matemáticos y lo real: en cuanto a las curvas, deben pasar por todos los pliegues del fenómeno, representarlo en todas sus fases, pero también abarcarlo en toda su amplitud y en todas sus partes.[44]

Conclusiones

Debido a esta posición crítica y a su interés por los temas de corte psicológico, se ha acusado a Halbwachs de cierta desconfianza hacia la utilización de estadísticas, pero su posición debe considerarse como una crítica saludable del mal uso que hacen los científicos sociales en los inicios de la Estadística matemática. Todavía hoy pueden encontrarse abundantes ejemplos del excesivo rigor aritmético y del olvido del estudio de lo concreto (en la línea de lo que Wright Mills denuncia como empirismo abstracto). Ideas como la superior complejidad de lo social frente a lo físico, o el hecho de que en la estadística, las cifras siempre se dejan combinar con cifras, todavía nos inducen a la reflexión.

Halbwachs y su maestro Simiand se anticiparon al estructuralismo, al sostener que con pocas repeticiones se pueden establecer "los mecanismos complejos" de los ciclos económicos. Ello conecta con el hecho social total de Marcel Mauss, o lo que también se ha llamado representatividad estructural: del mismo modo que en cada una de los individuos de un grupo biológico están contenidas las características de la especie, en el estudio de los hechos de naturaleza social pude procederse del mismo modo. Para Halbwachs, los hechos sociales están emparentados con la biología, y es del método biológico de donde extrae algunas de sus intuiciones sobre la investigación social.

En el apartado Science de la population et biologie correspondiente a Le point de vue du nombre,[45] sostenía con el apoyo de un texto de DARMOIS (Statistique et Aplications) [46], que

---

[41] En ella intervienen: M HUBER, director de la Estadística General de Francia con una ponencia titulada La Estadística: su historia,su organización, E.BOREL con La estadística. El instrumento matemático: el cálculo de probabilidades; A.PIGANIOL Y ED.ESMONIN: La estadística en Historia; M.HALBWACHS: La estadística en Sociología; H.LÉVY-BRUHL: La estadística y el derecho; P.VAN TIEGHEM: La estadística en Historia literaria; G.DARMOIS: La estadística en Sicología; G.TEISSIER: La estadística en Biología; M.BORN: La estadística en Física; P.LANGEVIN: Estadística y determinismo. Septième Semaine Internatione de Synthèse, 1935. Revue de Synthèse, 1944, Paris.

[42] ..tout comptage n'est pas une statistique. Revue de Synthèse (1944),p.115.

[43] Op. Cit. P.124.

[44] "Quand aux courbes, elles doivent subire tous les replis du phénomène, le représente dans toutes ses phases, mais aussi l' embrasser dans toute son étendue, et en toutes ses parties: c' est ainsi qu' on représentera les mouvements su salaire par plusieurs courbes, continués autant que possibles, juxtaposées, aussi nombreuses qu' il a de donnes correspondant à des groupes différents, agriculture, industrie, et diverses espèces et formes d' industrie, grandes villes, villes moyennes, petites villes. Cela exige un effort d' attention multiple, à la fois abstraite et concrète. Mais cette méthode empirique est la seule qui permette de rester en contacte aussi étroit que possible avec la réalité." Op. Cit, p.124.

[45] HALBWACHS M. Y SAUVY, A. (1936) Le pont de vue du nombre Édition critique sous la direction de MARIE JAISSON Y ÉRIC BRIAN (2005) Institut National d' Études Démographiques, p.229.

el método estadístico ha de pasar por tres fases: primero la presentación de las observaciones; segundo, su reducción, y tercero, la descripción, interpretación y explicación de regularidades estadísticas; en definitiva, lo que consideraba las fases del método científico. Su discurso más crítico va a girar en torno a un tema que le obsesiona como son las curvas teóricas, o lo que es igual, los modelos matemáticos, pues considera que desvían la atención de la producción de datos estadísticos. Sus advertencias en este sentido entrañan un amplio conocimiento de la sociedad y de las matemáticas sociales[47].

Resulta sorprendente la familiaridad de Halbwachs con los estadísticos y matemáticos más avanzados del momento. Maneja autores como Galton, Borel, Bowley, Yule, Darmois, etc., y está al tanto de los debates sobre la probabilidad. Tiene claras las diferencias entre la probabilidad "a priori"-de un modelo cerrado-, y la probabilidad "a posteriori", lo cual no quiere decir que no tenga reservas hacia la aplicación indiscriminada de ésta. Por ejemplo, nos dice, que cuando se quiere saber como se distribuye un fenómeno biológico o demográfico alrededor de una media, se parte de considerar la independencia de las causas de producción de los hechos como el fundamento del cálculo, pero la realidad es algo distinta porque los hombres de una misma época y de una misma región sufren en común la acción de las mismas fuerzas del mismo medio físico y social, es decir, los hechos sociales no son independientes como las bolas de una urna.

Pasado el período de euforia cuantitativista posterior a la segunda guerra mundial, la investigación social se ha abierto a otros enfoques y modos de investigación diferentes del quehacer estadístico. Halbwachs fue, además de un introductor de la estadística en las ciencias sociales, el precursor de nuevos métodos en la investigación social, y contrariamente a lo que se la ha afirmado, su interés por un conjunto de temas tan variado, no hace de él un científico disperso, sino un científico avanzado que supo conectar con ese rico pasado de la Ilustración y el Renacimiento europeos.

[46] DARMOIS publica Statistique Mathématique, el primer manual de estadística matemática francés, en 1928,
[47] si los matemáticos supieran hasta que punto los hechos sociales y su grado de complejidad son tan poco y mal conocidos, no intentarían encontrar con tanta presura la ley matemática de la población. Ibidem.

Bibliografía

AHEPE, (2002) Historia de la Probabilidad y de la Estadística. Ias Jornadas de Historia de la Estadística organizadas por la Asociación de Historia de la Estadística y de la Probabilidad de España. Madrid.

ALMENAR, S. (2002) Olegario Fernández Baños: de la geometría a la econometría., en FUENTES QUINTANA, E. Economía y economistas españoles, V6. La modernización de los estudios de economía. Fundación de las Cajas de Ahorros Confederadas. Círculo de Lectores., Madrid.

ARMATTE, M.(1992) Conjonctions, conjocture et conjecture. Les barometres économiques (1885-1930), Histoire et Mesure.

- (2002) El coeficiente de correlación y los economistas (1910-1940) en ARRIBAS, J. M Y BARBUT, M. Estadística y Sociedad, UNED.

ARRIBAS, J. M.(1994) Antecedentes de la Sociedad de Consumo en España : de la Dictadura de Primo de Rivera a la II República, Política y Sociedad, 16.

- (2002) ARRIBAS, J. M., Y BARBUT, M. (2002) Estadística y Sociedad UNED.
- (2002) ARRIBAS, J. M Y VALLEJOS A. Los orígenes de la Estadística Social en España: los trabajos de la Comisión y del Instituto de Reformas Sociales. En AHEP, Historia De la Probabilidad y de la Estadística. Madrid.

BLUM, A., Y MESPOULET, M. (2003) L'anarchie bureaucratique. Statistique et pouvoir sous Staline. La Découverte, Paris.

BOWLEY, A.L. (1906) Presidential Adress to the Economic Section of the British Association for the Advanced Sciences. Journal of the Royal Statistical Society.

- (1936): The application of sampling to economic and sociological probléms, Journal of the American Statistical Association, September,vol.31, nº 195. Traducción y presentación de José M. ARRIBAS, EMPIRIA, nº 5, 2002.

DESROSIERES, A. (1999) Del trabajo al consumo: evolución de los usos de las encuestas sobre el presupuesto de las familias. Anuario IEHS, 14. UNCPBA, Buenos Aires. Pp93-123.

- (2004) La política de los grandes números, Melusina, Barcelona.

FERNANDEZ BAÑOS, O. (1933) Aplicación del análisis estadístico a un problema económico. Economía Española, Oct.-Nov.-Dic., nº 10, 11,12.

- (1934) Sobre la correlación, medida de enlace directo o indirecto en los fenómenos económicos. Comunicación al IV Congreso de la Econometric Society, Stresa. Septiembre, Banco de España.
- (1941) Programa, concepto, método y fuentes de Estadística Matemática. Talleres Gráficos Marsiega, Madrid.

HALBWACHS, M. (1912) La théorie de l' homme moyen. Essai sur Quetelet et la statistiqu moral", Felix Alcan, Paris.

- (1913) La classse ouvrière et les niveaux de vie. Recherches sur la hiérarchie des besoins dans les societés industrielles contemporaines. Felix Alcan, Paris.
- (1930) Les causes du suicide. Librairie Félix Alcan, Paris.
- (1933) L'évolution des besoins dans les classes ouvrières. Librairie Félix Alcan, Paris.

- (1936) HALBWACHS M. ET SAUVY A., Le point de vue du nombre. Édition critique sous la direction de MARIE JAISSON ET ERIC BRIAN. (2005) Institut National d'Études Démografiques. Paris.
- (1937) Le point de vue sociologique. Edición electrónica de Jena-Marie Trembaly. « Les classiques des sciences sociales » WEB: http://www.uqac.uquebec.ca/zone30/Clasiques des_sciences_sociales/index.html Jean-Marie Tremblay
- (1944) La Statistique. Ses applications. Les problèmes qu'elles soulèvent. Septième Semaine Internationale de Synthèse, 1935.

HORMIGÓN M. (1998) Las matemáticas en España en el primer tercio del siglo XX, en Sánchez Ron JM, Ciencia y Sociedad en España, Ediciones El arquero, CSIC.

JAISSON M., BOUDELOT. C. (2007) Maurice Halbwachs, sociologue retrouvé. Editions Rue d'Ulm, Paris.

MACKENZIE, D A, (1981) Statistics in Britain 1865-1930 Edinburgh University Press

MARTIN O. Raison statistique et raison sociologique chez Maurice Halbwachs, en Revue d' Histoire des Sciences Humaines (1999)

MIGUEL, A. DE (1924) Metodología Estadística. Fundamentos de Estadística matemática, Madrid.

POLLACK, M. (1979) Paul Lazarsfeld, fundador de una multinacional científica. En Álvarez Uría, F. y Varela, J (eds.) Sociología Crítica, Madrid, La Piqueta, 1986. pp-37-82.

REVUE D'HISTOIRE DES SCIENCES HUMAINES (1999), nº 1. Maurice Halbwachs et les sciences humaines de son temps. Septentrion, Presses Universitaires.

REY PASTOR (1988) Selecta, Fundación Banco Exterior.

SHAPIN S.(1998) La Révolution Scientifique, Flammarion 1998.

TOOZE J.A. (2001) Statistics and the German State, 1900-1945. The Making of Modern Economic Knowledge. Cambridge, University Press.

Web:, http://www.uqac.uquebec.ca/zone30/Clasiques des_sciences_sociales/index.html Jean-Marie Tremblay

Capítulo 6

Homenaje al profesor Dr. D. Francisco Azorín Poch

ANTONIO FRANCO RODRÍGUEZ-LÁZARO
Mª CARMEN ESCRIBANO RÓDENAS
ANDRÉS GUTIÉRREZ GÓMEZ
Universidad CEU San Pablo

Introducción

Siempre es una satisfacción rendir homenaje a personas ilustres que han dejado una profunda huella en algún campo de actividad o del saber, y lo es mucho más cuando se trata de una figura señera en diferentes frentes como son el académico, el profesional, el cultural, y el humanista, y que en el caso que nos ocupa está adornada, además, con unas cualidades personales como la tolerancia, la modestia, la generosidad y la bondad. Nos estamos refiriendo al profesor Dr. D. Francisco Azorín Poch académico numerario de la Real Academia de Ciencias Físicas, Exactas y Naturales, Estadístico facultativo, y Catedrático de Universidad.

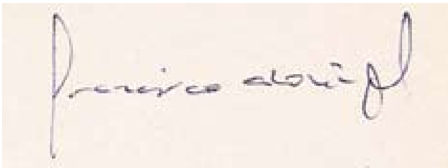
Fue un hombre de gran lucidez intelectual y dignidad moral, de gran competencia científica y pedagógica, siempre actualizado, que participó en labores creadoras, con auténtica vocación y talante universitarios, de gran categoría científica reconocida nacional e internacionalmente.

Fue un auténtico maestro, del saber y de los valores, que él encarnó y que sin ser consciente de ello inculcó a sus discípulos en los distintos centros universitarios en los que ejerció como docente, no sólo en sus clases de licenciatura, sino en los cursos de doctorado y en la dirección de trabajos de investigación y de tesis doctorales, que él no sólo sugería y dirigía, sino que participaba con una dedicación constante, que llegaba a la corrección de pruebas de imprenta.

Como un erudito que era de terminología del ISI era partidario de la escritura y del discurso precisos, buscando siempre el vocablo adecuado y ahí se reflejaba no sólo su vasta cultura, sino sus profundos conocimientos de idiomas como latín, griego, español, inglés, francés, alemán, ruso, árabe y esperanto[1], en el que impartió conferencias en las que llamó la atención de sus asistentes por su precisión y entusiasmo.

Llegados a este punto cabe decir aquí que era un gran aficionado al estudio de la literatura, de las religiones y de las ciencias naturales, y entre éstas de la Botánica, la Biología y la Ornitología. Concretamente, durante su estancia en Chile trabó conocimiento con un famoso entomólogo que quedó admirado por los conocimientos científicos del profesor Azorín sobre insectos y aves, y comentó la exquisitez de los dibujos que hacía en su cuaderno de campo después de observar con sus prismáticos especiales los especímenes que encontraba.

Datos Biográficos



Francisco Azorín Poch nace en Málaga el día 2 de Julio de 1914. Su madre Carmen Poch es natural de Málaga, y su padre Francisco Azorín Izquierdo es natural de Teruel.

El arquitecto Francisco Azorín Izquierdo nació el 12 de septiembre de 1885 en Monforte de Teruel y murió en México en 1976. Perteneció al PSOE (Partido Socialista Obrero Español), ocupando el cargo de concejal en Córdoba y posteriormente el de diputado en Cortes antes de la Guerra Civil española. Como arquitecto, entre otras obras construyó la Casa del Pueblo de Madrid. Fue un personaje muy activo en el movimiento a favor del idioma esperanto por el que empezó a interesarse en 1910[2], e intentó que su hijo Francisco aprendiese, con el esperantista Joseph Berger Knöll que residía en Córdoba en aquella época. Escribió un diccionario ilustrado de esperanto (*Ilustrita Vortaro de Esperanto*[3]) y otro de términos de arquitectura, con el esperanto como idioma base, y con la traducción de los términos a numerosos idiomas, *Universala Terminologio de la Arkitekturo*, y fue elegido representante del movimiento esperantista español en la Asociación Mundial de Esperanto. Fue cónsul de España en la localidad francesa de Toulouse durante la guerra civil, y tras el final de ésta tuvo que exiliarse a México, donde falleció.

Francisco Azorín Poch inicia sus estudios de Bachillerato en 1924, en el Instituto Nacional de Enseñanza Media de Córdoba, donde reside su familia en este momento. En este instituto fue alumno del profesor D. José Manuel Camacho[4]. En esta época Azorín ya llevaba años expresando que deseaba llegar a ser catedrático de Historia Natural.

En 1931 inicia sus estudios universitarios en la Universidad Central de Madrid, simultaneando la carrera técnica de Aparejadores con la de CC. Exactas[5]. Aunque se ha publicado por el INE[6] que estudió CC. Químicas no parece comprobado. En esta época de estudiante vive en la Residencia de Estudiantes de la calle Pinar número 21 de Madrid, y a la

[1] Esta afición por el esperanto se la inculcó su padre D. Francisco Azorín Izquierdo.
[2] En 1910 asistió a un congreso del PSOE celebrado en Copenhague y se entusiasmó por este idioma.
[3] De esta obra sólo se publicó la primera parte, y la segunda quedó como manuscrito no publicado.
[4] Este profesor le dedicó un artículo en 1930, con ocasión de habérsele concedido un premio literario, donde explica que Azorín, que "siempre estaba ávido de obtener nuevos y diversos conocimientos, posee un espíritu ordenado y serio…. Y anotó el hecho, quizá pensando que mañana pudiera servirle para alguna estadística".
[5] El profesor Ríos dirá de él que en esta época "no tenía clara la senda de su vocación".
[6] Véase la cronología del profesor Azorín, publicada en la contraportada del libro *Selección de escritos estadísticos del profesor Azorín Poch aparecidos en publicaciones especializadas*, que publicó el INE en 1984, con motivo de un homenaje que se le rindió por su jubilación.

que Juan Ramón Jiménez llamó la "Colina de los Chopos". Al igual que muchos de los mejores intelectuales, poetas y científicos españoles de la época, adquiere en ella el hábito de trabajar con espíritu y metodología universitaria, además de participar en actividades humanísticas literarias o plásticas.

El 18 de Julio de 1936 Francisco Azorín estaba con sus padres en la estación de ferrocarril de Córdoba a la espera de tomar el tren que le conduciría a Málaga para pasar las vacaciones, cuando un compañero de su padre vino a avisar del levantamiento militar. Su padre decidió continuar con los planes establecidos, llevar a su familia a Málaga y tomar el primer tren de vuelta a Córdoba. Sin embargo no pudo hacerlo, teniendo que exiliarse, de Málaga pasó a Barcelona y de allí a la ciudad francesa de Toulouse, donde sería nombrado Cónsul de la República Española. En 1939 toda la familia de Francisco Azorín Poch embarcaría en el vapor "Ipanema" rumbo al Puerto de Veracruz (México)[7], él fue el único que permaneció en España y estuvo algún tiempo en la cárcel por su participación en la Guerra Civil como teniente de artillería del ejército de la República.

Según el propio Azorín fue la lectura de un libro de Borel y la sensación de estar ante una acertada representación del razonamiento matemático al escuchar al profesor D. Olegario Fernández Baños en el viejo caserón de San Bernardo, un día que explicaba la curva de los precios del boniato, las experiencias que motivaron su dedicación a la Estadística. Dándose cuenta con ellas de las diferentes repercusiones y reflexiones que pueden derivarse del lenguaje matemático, al comparar el desarrollo metodológico de las otras asignaturas de la carrera de Exactas con la apertura de miras y la tolerancia con la que las generalizaciones matemáticas intentan dar respuesta a experiencias, motivaciones y conductas humanas.

Son sus dotes de observación y sistematización, así como la disposición a incrementar de forma constante sus capacidades y conocimientos, con el propósito de estar en situación ventajosa ante los diferentes interrogantes y obstáculos a los que se fue enfrentando a lo largo de su vida, los que le llevan a dedicar su actividad profesional a la Estadística.

En 1943, una vez ha finalizado la guerra civil, termina la Licenciatura en Ciencias Exactas por la Universidad Central de Madrid y al año siguiente ingresa con el número uno de su promoción en el cuerpo de Estadísticos Facultativos y obtiene su primer destino en la Delegación de Segovia. Simultánea esta ocupación con la plaza de ayudante de cátedra en la universidad de Madrid[8], primero bajo la dirección de Olegario Fernández Baños y posteriormente de Sixto Ríos. Fue compañero de cátedra de Enrique Cansado[9] que ocupaba el cargo de auxiliar temporal. Posteriormente, pasa a ser becario del Instituto San José de Calasanz de Pedagogía y del Instituto de Ampliación de Estudios e Investigación. Participa en el primer seminario que se imparte en España sobre muestreo estadístico, siendo el INE el organizador de las conferencias. Interviniendo ulteriormente en un seminario de este mismo tema en la Universidad de Estocolmo.

En 1947 contrae matrimonio con Carmen Mínguez Blas, licenciada también en Ciencias Exactas.

[7] Ángel Azorín Poch, hermano de Francisco, fue arquitecto que ejerció en México

[8] El primer profesor de esta cátedra fue Esteban Terradas en el curso 1931-32, y el primer catedrático por oposición Olegario Fernández Baños que la ocupó desde 1934 hasta su muerte en 1945. El sucesor fue Sixto Ríos, del que fue compañero y que posteriormente daría la contestación a su discurso de ingreso en la Real Academia de Ciencias.

[9] El profesor Enrique Cansado fue presidente electo del I.S.I. (Instituto Internacional de Estadística).

Dentro del Departamento de Estadística[10] del Instituto de Matemáticas del Consejo Superior de Investigaciones Científicas (C.S.I.C.) bajo la dirección del profesor Sixto Ríos, se empezaron a desarrollar cursos y seminarios impartidos por profesores extranjeros invitados, lo que le permitió conocer personalmente a Frechet, Mahalanobis, Wold, Cramer y Wishart. El primer profesor invitado, con la colaboración del Instituto de Investigaciones Agronómicas, fue el Dr. J. Wishart en 1947, que inició un grupo de investigadores que trabajarían en Estadística Aplicada al diseño de experimentos.

El profesor H. Wold, de la universidad sueca de Upsala, también impartió un seminario e impulsó el desarrollo de la Estadística en España, siendo una de sus consecuencias la creación de la revista "Trabajos de Estadística" que comenzó a publicarse en 1950 con una regularidad de tres números anuales. En esta revista Francisco Azorín empezó a colaborar asiduamente, a partir del primer número de la misma.

Su espíritu abierto y lleno de sana curiosidad le llevó a realizar una serie de viajes por el mundo, comenzando por el que hizo a Inglaterra en el curso 1951-52 con la finalidad de completar su formación académica en el laboratorio de Estadística de la Universidad de Cambridge, en calidad de investigador visitante becado por el Consejo Británico; lo que le permitió disponer de la experiencia y enseñanza directa de investigadores de gran preeminencia científica, entre los que estaban Wishart, Lindley, Cox y Anscombe.

En 1953 obtiene el doctorado en Ciencias Exactas por la Universidad de Madrid, con premio extraordinario, la tesis se titula la "Distribución t no central".

La década de los años 50 tuvo muchas novedades en la disciplina de la Estadística. Se creó la Escuela de Estadística[11], y el Instituto de Investigaciones Estadísticas del CSIC. En ambos colaboró Azorín.

Posteriormente, se desplazó de forma temporal a América Latina donde pasó a ocupar puestos relevantes, en unos casos como profesor y en otros como experto estadístico. En 1955 colaboró en la creación de la Escuela de Estadística que bajo el auspicio de la UNESCO se estableció en la Universidad Central de Venezuela, llegando a ser el Jefe del grupo de expertos que se encargó de organizar su puesta en funcionamiento y donde después permaneció impartiendo docencia de Estadística. Es durante su estancia en esta universidad cuando publica la primera edición de su Curso de Muestreo y Aplicaciones, obra de referencia básica en esta materia que le convirtió en una autoridad a nivel internacional.

En 1958 es nombrado Jefe de la Sección de Coordinación del Instituto Nacional de Estadística español.

Fue un hombre que se sentía ciudadano del mundo pero sin perder sus connotaciones españolas. Las experiencias y avatares vividos en los viajes le permitieron desarrollar su capacidad para relacionarse con matemáticos de otros países y concienciarse de la importancia de aprender otros idiomas. Posteriormente, esta decisión le iba a permitir una fluida captación de las nuevas teorías estadísticas y el poder publicar los resultados de sus investigaciones en las revistas científicas de otros países.

Pero no solamente llegó a dominar varios idiomas entre los que estaban el inglés, el alemán y el ruso, además del árabe, el latín y el griego; también se le podía considerar un

[10] Posteriormente sería el Instituto de Investigaciones Estadísticas (I.I.E.) del C.S.I.C.

[11] Véase ESCRIBANO RÓDENAS, M. C.; BUSTO CABALLERO, A. I. (2002): "La creación en España de la primera Escuela de Estadística" en *Historia de la Probabilidad y de la Estadística*. A.H.E.P.E. Madrid.

experto en Botánica, Ornitología y poseía grandes dotes para el dibujo. A todas estas aptitudes hay que añadir que era un gran cultivador de la amistad de sus amigos, que unido al hecho de que poseía una extensa cultura, le permitía hacer amenas las conversaciones sobre literatura o religión, así como las frecuentes visitas a museos, asistencia a exposiciones o contemplar un simple monumento.

Obtiene por oposición la cátedra de Estadística Matemática y Cálculo de Probabilidades de la Universidad de Santiago de Compostela en 1961, un año después es nombrado asesor regional de muestreo y jefe del Centro de Proyecciones económicas, y algunos años más tarde Director de la División de Estadística de la CEPAL de Naciones Unidas, con residencia en Santiago de Chile. Su vida se acelera y se hace más compleja a partir de este momento.

No es fácil tomar la vida como un reto, compaginando el desempeño de cargos en instituciones y a la vez desarrollar una incesante actividad investigadora. Por eso es digna de elogiar su destacada actuación en el ámbito internacional como profesor de muestreo bajo los auspicios de la UNESCO y a partir de 1965 es elegido miembro de número del Instituto Internacional de Estadística (International Satistical Institute I.S.I.). Al cabo de unos años vuelve a España y en 1974 toma posesión de la cátedra de Estadística Matemática de la Universidad Autónoma de Madrid, siendo desde entonces el Director del Departamento de Estadística de la Facultad de Ciencias de la citada universidad, hasta Diciembre de 1982.

Le nombran Presidente del Instituto Nacional de Estadística, INE en 1977 con la misión de lograr que las estadísticas oficiales tengan una mayor repercusión en las expectativas de las empresas cuando diseñan sus planes de futuro y en la aplicación de la política económica en nuestro país. Hasta 1982 permanece en la presidencia del INE, lo que le confiere la obligación de asumir la representación del Instituto en diferentes reuniones internacionales, en las que logró el respeto y la consideración de la comunidad estadística internacional hacia la labor desarrollada por los estadísticos españoles.

En 1980 es nombrado miembro de la Comisión de Terminología de la IASS (Internacional Association of Survey Statisticians), y al año siguiente ingresa como miembro en la Real Academia de Ciencias Exactas, Físicas y Naturales[12]. Azorín en su discurso de ingreso efectuó un excelente resumen de la teoría de los conjuntos borrosos o difusos que en ese momento se encontraba en sus inicios. Sixto Ríos fue el encargado de realizar el discurso de réplica al efectuado por Azorín, en él resalta los méritos científicos y profesionales, así como las cualidades humanas de bondad, modestia, simpatía y tolerancia de un hombre que supo distinguir durante toda su vida lo fundamental de lo accesorio, orientando su curiosidad científica y capacidad de análisis a la obtención de conocimientos que fueran útiles para los demás, sin abandonar en ningún momento la alegría de vivir.

Como miembro del ISI, y en colaboración con el Embajador de España en Manila (Filipinas), en una de las reuniones celebradas, propone que Madrid sea la sede del 44º período de sesiones del ISI, que tendría lugar en 1983 y para el cual fue designado Presidente del Comité Ejecutivo de la Comisión Organizadora Nacional.

En la madurez profesional continuó dando numerosa conferencias en universidades españolas y extranjeras, siguió impartiendo clases como catedrático en la Universidad Autónoma de Madrid, a la vez que participaba en las actividades de la Real Academia de Ciencias Exactas, Físicas y Naturales.

[12] Su discurso de ingreso se titula "Conjuntos borrosos, estadística y probabilidad", y fue contestado por el académico, amigo y compañero, el profesor Sixto Ríos García, quien afirmó que "Azorín nació para ser estadístico".

La jubilación como estadístico facultativo y profesor universitario llegó en 1984 y con tal motivo se le rinden sendos homenajes; por una parte, del Instituto Nacional de Estadística, y por otro lado, el Departamento de Estadística y División de Matemáticas de la Universidad Autónoma de Madrid.

A propuesta de la Facultad de Ciencias Económicas y Empresariales de la Universidad de Málaga fue investido Doctor Honoris Causa en 1985.

El profesor Azorín Poch siguió desarrollando su actividad científica en la Academia de Ciencias, impartió conferencias y asistió a Congresos permaneciendo en activo ininterrumpidamente hasta que en 1989 dictó su última lección magistral como afirmó el profesor Sánchez-Crespo:

> *"Su muerte fue su última lección magistral, que no tuvo necesidad de preparar. Le fue suficiente con hacer lo cotidiano. Sin estruendo, con sigilo, casi de puntillas, sabiendo distinguir lo fundamental de lo accesorio, dedicando su atención y sus sonrisas a todos los demás, pacífico, pero venciendo como guerrero heroico los zarpazos de su terrible enfermedad, y, siguiendo una línea de conducta hasta el último momento. Cuando apenas le quedaban doce horas de vida y a pesar del gota a gota y la respiración asistida, me comentaba un determinado libro"*[13].

Rasgos relevantes de la obra científica del profesor Azorín

Los artículos, libros, conferencias y demás formas utilizadas por Azorín para divulgar sus planteamientos científicos reflejan la gran pasión que sentía por ahondar en todo tipo de conocimientos, trabajando y aportando nuevas ideas en todos los temas punteros de su época desde que veían la luz en las publicaciones de carácter científico.

El rigor y la profundidad del razonamiento matemático de Azorín se ponen de manifiesto en las "Conferencias de Preparación Matemática y Estadística" publicadas por el INE en 1950, al especificar con gran detalle los conocimientos que había en ese momento sobre la medida de conjuntos y las integrales de Lebesgue-Stieljes, conceptos básicos necesarios para estudiar las distribuciones de probabilidad, añadiendo aspectos que incrementaban la precisión de las explicaciones y facilitaban su comprensión.

Es usual que los intelectuales con interés en profundizar en las cosas hagan incursiones en diferentes campos del conocimiento logrando desarrollar con éxito todos los aspectos investigados en ellos. Por eso, no debemos extrañarnos que el profesor Azorín haya realizado importantes aportaciones en poblaciones finitas, análisis multivariable, conjuntos borrosos, etc.

Las investigaciones realizadas por Azorín versan sobre las dificultades que surgen en la Estadística Inferencial a la hora de tener que elegir los procesos de estimación adecuados, al evaluar y clasificar los estimadores, al seleccionar las características deseables de estos estimadores y de los datos observados, especialmente en el Muestreo en Poblaciones Finitas y en el diseño y análisis de las encuestas por muestreo. Destacando sus trabajos en muestreo espacial, de tapices o mosaico.

[13] *En recuerdo de Francisco Azorín Poch*, por J.L. SÁNCHEZ-CRESPO, en "Estadística Española", vol. 31, núm. 120, 1989, Pág. 131.

La Universidad Central de Venezuela publicó su libro "Curso de Muestreo y Aplicaciones" que recoge las lecciones impartidas en esa Universidad sobre muestreo en poblaciones finitas. Años después revisa su contenido durante la estancia que tuvo como Catedrático de Estadística y Cálculo de Probabilidades en la Universidad de Santiago de Compostela, edición que publica el INE. La tercera revisión fue editada en Madrid por Aguilar, pero realizó su redacción durante el tiempo que vivió en Santiago de Chile, donde llegó en calidad de Asesor Regional de Muestreo de las Naciones Unidas y después como Director de la División de Estadística de la Comisión Económica para América Latina de la ONU. El esfuerzo y el tiempo dedicado a desarrollar el Muestreo en Poblaciones Finitas le llevó a ser una figura mundialmente reconocida en este campo de la Estadística.

Este fue uno de los temas que más ayudó a divulgar entre los estadísticos españoles y latinoamericanos, llegando a concienciar a sus coetáneos de la importancia de comenzar a aplicar cuanto antes un método de análisis que era mucho más realista al realizar investigaciones concretas que el considerar que las poblaciones son infinitas. Su nombramiento como Presidente del INE y el ser profesor de diferentes universidades facilitaron el proceso de difusión del Muestreo de Poblaciones Finitas.

En la literatura científica existen los denominados "libros clásicos" que son aquellas referencias universales de las que se suele extraer la comida científica, el maná para la adquisición de conocimientos, como base y referencia de sus trabajos y actividades el profesor Azorín es autor de un clásico Curso de muestreo y aplicaciones, cuya tercera revisión editó en Madrid la editorial Aguilar.

> *"La preparación del libro se hizo en tres ciudades que dejaron en él un recuerdo inolvidable, y a los "amigos y colegas caraqueños, santiagueses y santiaguinos" dedicó su agradecimiento y los tres santiagos estuvieron siempre en su memoria, pues decía "...en Caracas, esto es, Santiago de León de Caracas, como fue su nombre histórico original""*[14].

Le apasionaba el tema de la reconstrucción de un patrón o mosaico a partir de una muestra extendida al espacio euclídeo que lo contiene. De igual forma, la Taxonomía Matemática ocupaba un lugar preferente en sus investigaciones, porque incluía aplicaciones de la partición de conjuntos y de la idea de distancia o proximidad, conceptos que desafiaban su capacidad de relacionar el mundo empírico con el mundo conceptual. Esta manera de ser hizo que se convirtiera en un experto en Análisis Discriminante y Análisis Cluster, ya que ambos están muy relacionados con el Reconocimiento de Patrones. La Universidad de Manchester le reconoció ese prestigio al encargarle que redactara el capítulo sobre Análisis Discriminante para la Systems and Control Enciclopedia.

En 1979 el INE le publicó la primera obra en español sobre Conjuntos Borrosos "Algunas aplicaciones de los conjuntos borrosos a la Estadística" y en el acto de su ingreso como miembro de la Real Academia de Ciencias Exactas, Físicas y Naturales leyó un discurso que llevaba por título "Conjuntos Borrosos, Estadística y Probabilidad".

Es difícil incorporar grandes dosis de imaginación en la investigación de temas estadísticos y ver las posibilidades de aplicación futura de los nuevos planteamientos metodológicos o de las técnicas de análisis que van surgiendo a lo largo del tiempo, todo depende de la decisión que tome el científico en un instante decisivo en el que debe elegir si

[14] *En recuerdo de Francisco Azorín Poch*, por J.L. SÁNCHEZ-CRESPO, en "Estadística Española", vol. 31, núm. 120, 1989, Pág. 132.

dedicará una parte importante de su vida a la búsqueda de planteamientos nuevos o de soluciones mediante el estudio de un aspecto concreto de la realidad.

Conclusiones

Todo lo reseñado no hubiera sido posible si el profesor Azorín no hubiese tenido a su lado una persona que se identificó con él, le amó y le acompañó hasta el último momento, proporcionándole esa latente felicidad que sólo pueden proporcionar las personas excelsas, nos referimos a su mujer Carmen que le dio como fruto de su matrimonio un hijo, Ernesto Azorín Mínguez, prestigioso profesional de la Estadística lo que corrobora el destacado puesto que ocupa en EUROSTAT (la Oficina de Estadística de la Comisión Europea) en el ámbito de la cooperación estadística internacional, con sede en Luxemburgo, y que es a su vez padre de los dos nietos del profesor Azorín.

Y como Azorín, además de todo lo dicho era una buena persona terminamos esta ponencia con una frase de Concepción Arenal que refleja nuestros sentimientos acerca de la persona que nos ocupa:

"El mejor homenaje que puede tributarse a las personas buenas es imitarlas".

Catálogo de publicaciones del profesor Francisco Azorín Poch

Conferencias de Preparación Matemática y Estadística, publicadas por el INE en 1950.
Product Sums and Modulus Sums of H. Wold´s Normal Deviates (Sumas de productos y sumas de módulos), en colaboración con H. Wold, "Trabajos de Estadística". Vol. I. Cuaderno I, 1950. Págs. 5-28.

Sobre el apuntamiento y su medida (On peakedness and its measure), en "Trabajos de Estadística". Vol. I. Cuaderno III, 1950. Págs. 263-272.

Los funcionarios científicos y los estadísticos oficiales en Gran Bretaña, en 5º Suplemento al Boletín del INE, año 14. Madrid.

Notas sobre Taxonomía y Estadística, en "Trabajos de Estadística". Vol. XIII. Cuaderno III, 1952. Págs. 249-263.

Aportación de experiencias a los trabajos del Instituto Nacional de Estadística (1er centenario de la Estadística Española), I.N.E., Madrid, 1956.

Algunos problemas estadísticos en la construcción de escalas de consumo, en "Trabajos de Estadística". Vol. X. Cuaderno II, 1959. Págs. 63-73.

Consideraciones estadísticas sobre la función de Cobb-Douglas, en "Estadística Española", nº 4, Julio-Septiembre 1959. Págs. 19-29.

Nota sobre la estimación de varianzas por el método de grupos aleatorios, en Revista de la Academia de Ciencias Exactas, Fisico-Químicas y Naturales de Zaragoza, Serie 2ª, tomo XV, fascículo 2º. Zaragoza, 1960.

The teaching of statistics in Spain, en "The American statistician", revista de la American Statistical Association, Vol. 14, nº 4, Octubre 1960. Págs. 23-26.

Sobre el muestreo en bola de nieve, "Estadística Española". Nº 14, Enero-Marzo, 1962. Págs. 13-21.

Sobre la asociación en sucesiones de atributos, en "Trabajos de Estadística". Vol. XIV. Cuadernos I y II, 1963.

Algunas aplicaciones de los criterios de espaciamiento óptimo, en "Trabajos de Estadística y de Investigación Operativa". Vol. XIX. Cuaderno III, 1968. Págs. 3-49.

Nota sobre aplicación de algunos criterios de valor óptimo, en "Estadística Española", nº 46, Enero-Marzo, 1970. Págs. 5-18.

Comparación de supuestos y diseños en la estimación de patrones por muestreo sistemático pluridimensional, en "Trabajos de Estadística e Investigación Operativa". Vol. XXI, Cuaderno III. Madrid, 1970.

Los Estadísticos: funciones, profesión y perspectivas, en "Estadística Española", número conmemorativo de los XXV años de la Ley de Estadística, 1971.

Reconstitución de Patrones por muestreo sistemático espacial, en "Estadística Española", nº 53, Octubre-Diciembre 1971. Págs. 5-43.

Nota sobre modelos estadísticos en problemas de comportamiento, en "Estadísticas y Actuariado", Revista del Colegio de Estadísticos y Actuarios de Venezuela (CEAV), nº 8, Diciembre, 1971. Págs. 80-90.

Identificación de sistemas mediante aproximación diferencial segmental, Santiago de Chile, 1972.

Trade-Offs in Taxonomy, en Boletín del Instituto Internacional de Estadística (I.S.I.), Proc. 39 Session, Viena, Austria, 1973.

Data Banks- Problems of Classification and Computerization, 2nd. Conference of the IARIW, Rio de Janeiro, Brasil, 1974.

Estudio sobre la asignación de recursos para el mejoramiento de las fuentes de estadísticas demográficas en los países de América Latina, en colaboración con C. Cavallini, borrador para discusión, Naciones Unidas, CEPAL, Santiago, Chile, 1974.

On statistical models as a subject of classification and resource allocation, en Boletín del Instituto Internacional de Estadística (I.S.I.), Varsovia, Polonia, 1975. Págs. 42-46.

Criterios taxonómicos y de asignación aplicados a modelos, en "Trabajos de Estadística", Vol. XXVI, Cuadernos I, II y III, 1975.

Estadística y Taxonomía matemática, en "Estadística Española", nº 70 y 71, Enero-Junio 1976. Págs. 89-99.

Conjuntos no nítidos y modelación laxa en taxonomía matemática, Coloquio Internacional de Estadística e Investigación Operativa, Consejo Superior de Investigaciones Científicas. Madrid, 1976.

Statistique Mathematique. Sur l´entropie taxonomique des classifications, en colaboración con J. Garrido, en « C.R. Acd. Sc. Paris », tomo 284, 4 abril 1977, Serie A, Págs. 819-822.

On fuzzy simples from fuzzy populations, en Boletín del I.S.I., Proc. 41 Session, New Delhi. India, 1977.

Muestreo borroso de poblaciones borrosas, en "Estadística Española", nº 78 y 79. Madrid, 1978.

Muestreo de poblaciones borrosas, reunión L.A. de Muestreo. Universidad Autónoma Metropolitana de México, D.F., 1978.

Información estadística, caracterización y transacciones, presentación a la Sesión Inaugural de la X reunión anual de Investigación Operativa, Estadística e Informática en Madrid, Noviembre de 1977, y publicado en "Estadística Española", nº 80 y 81, Julio-Diciembre 1978. Págs. 17-24

Algunas experiencias de muestreo en el Instituto Nacional de Estadística de España, en colaboración con J.L. Sánchez-Crespo reunión latinoamericana de Muestreo. Universidad Autónoma Metropolitana, México, 1978.

Some methodological problems in large scale simple surveys (lsss) for developing countries, en colaboración con J.L. Sánchez-Crespo, en Boletín del I.S.I., Vol XLVIII, Manila, 1979.

Algunas aplicaciones de los conjuntos borrosos a la Estadística. I.N.E., Madrid, 1979.

Estadística Sistémica, en "Estadística Española", nº 82 y 83, Enero-Junio 1979. Págs. 9-15. Conferencia impartida en el Auditórium del Palacio de las Academias de Caracas (Venezuela), el 13 de Noviembre de 1978.

Conceptos y criterios de evaluación, Ciclo de conferencias de Investigación Operativa, publicado por la Real Academia de Ciencias Exactas, Físicas y Naturales. Madrid, 1980. Págs. 31-52.

Contribution to the balancing of various desiderata attributes, presentado en la Sesión plenaria de la Conferencia de Estadísticos Europeos/424/Add.3, celebrada en Ginebra, Suiza, en Junio de 1980.

Contribución al equilibrio de varias características deseables, en colaboración con J.L. Sánchez-Crespo, en "Estadística Española", nº 87, Abril- Junio 1980. Págs. 7-13.

Algunas reflexiones en torno a la información estadística y la noción de conexión, en "Estadística Española", nº 88, Julio-Septiembre 1980. Págs. 7-15. Discurso de apertura del II curso Intensivo de Muestreo con Aplicaciones a Encuestas de Hogares, realizado en el I.N.E., Madrid, Abril de 1980.

Algunos aspectos metodológicos en las encuestas continuas a gran escala, en colaboración con J.L. Sánchez-Crespo, Experiencias recientes del I.N.E. , Sao Paulo, Brasil, 1980.

Ventajas e inconveniente en la sustitución de las investigaciones estadísticas exhaustivas por encuestas muestrales, en colaboración con J.L. Sánchez-Crespo, en "Estadística Española", nº 90, Enero-Marzo 1981. Págs. 5-9.

Conjuntos Borrosos, Estadística y Probabilidad, discurso leído en el acto de su recepción en la Real Academia de Ciencias Exactas, Físicas y Naturales, el 2 de Diciembre de 1981.

Borrosidad y robustez, publicado por la Facultad de Ciencias y el Dpto. de Estadística Matemática de la Universidad Autónoma de Madrid, 1982. Págs. 573-583.

Field supervisión zones in accordance with the correlated component of the response variante, en colaboración con J.L. Sánchez-Crespo, presentado en el Seminario Naciones Unidas, Conferencia de Estadísticos Europeos, Moscú, 21-25 Septiembre 1981.

Zonas de supervisión de acuerdo con el componente correlacionado de la varianza de respuesta, en colaboración con J.L. Sánchez-Crespo, en “Estadística Española”, nº 94, Madrid, 1982. Págs. 9-19.

Estadística, Ciencia y Sociedad, publicado por Revista de la Academia de Ciencias Matemáticas, Físico-Químicas y Naturales de Granada, Mayo, 1982. Vol. I. Págs. 6-34.

El futuro de la Estadística, Actas de la XII reunión nacional de Investigación Operativa, Estadística e Informática, Jaca, 1980. Publicado en “Cuadernos de Bioestadística” de la Facultad de Medicina y Facultad de Ciencias de la Universidad de Zaragoza. Vol. I, nº extra, 1983.

Muestreo de Poblaciones Finitas, en “Estadística Española”, Págs. 34-41.

La inferencia estadística, en “Historia de la Ciencia Estadística”. Real Academia de Ciencias Exactas, Físicas y Naturales. Ciclo de conferencias. Madrid, 1989. Págs. 45-64.

Libros

Algunas aplicaciones de los Conjuntos Borrosos a la Estadística. Instituto Nacional de Estadística. Madrid, 1979.

Curso de Muestreo y Aplicaciones. Aguilar. Madrid, 1969.

Curso de Muestreo y aplicaciones. Instituto Nacional de Estadística. Madrid, 1982.

Métodos y aplicaciones del muestreo, en colaboración con J.L. Sánchez-Crespo. Alianza Universidad Textos. Madrid, 1986.

Bibliografía

ESCRIBANO RÓDENAS, M. C.; BUSTO CABALLERO, A. I. (2002): “La creación en España de la primera Escuela de Estadística” en *Historia de la Probabilidad y de la Estadística*. A.H.E.P.E. Madrid. Págs. 205 – 220.

INSTITUTO NACIONAL DE ESTADÍSTICA (1984): Selección de Escritos Estadísticos del profesor Francisco Azorín Poch. Madrid.

REAL ACADEMIA DE CIENCIAS EXACTAS, FÍSICAS Y NATURALES (1989): *Historia de la Ciencia Estadística.* Ciclo de conferencias. Madrid, 1989.

SANCHEZ-CRESPO, J.L. (1989): “En recuerdo de Francisco Azorín Poch”, *Estadística Española*, 31 (120) Págs. 131-134.

Capítulo 7

Evolución de la enseñanza de la estadística en España a lo largo del siglo XIX

ANA I. BUSTO CABALLERO
Universidad Complutense
Mª CARMEN ESCRIBANO RÓDENAS
Universidad CEU San Pablo

Introducción

A finales del siglo XVIII y principios del XIX los gobiernos empiezan a tomar conciencia de lo importante que es para una nación la educación de sus ciudadanos.

Sin embargo, la complicada política española de estos momentos hace que los esfuerzos de los diferentes gobiernos por mejorar la instrucción pública fracasen.

El Plan Caballero, de 1807, no llega a implantarse por la contienda con los franceses; las mejoras en la educación que prometía la Constitución de las Cortes de Cádiz, con el informe Quintana (1813) como consecuencia, no pudieron llevarse a cabo más que en el Trienio Constitucional (1820-1823) por las reacciones absolutistas de 1814 y 1823.

Es por eso que la primera configuración del sistema educativo español no se empieza a producir hasta después de la muerte de Fernando VII (1833), cuando vuelven los liberales exiliados y se inicia la revolución liberal.

A partir de entonces se van sucediendo cambios importantes en la educación española: se crean las escuelas de párvulos, las escuelas normales, los institutos de enseñanza secundaria; se legisla minuciosamente todo el sistema educativo, se editan numerosos libros de texto, se leen las primeras tesis doctorales...

En resumen, el siglo XIX caracterizado por la inestabilidad política es testigo del nacimiento del sistema educativo español.

Las primeras cátedras de Estadística españolas

Ya el Informe Quintana (1813) y el proyecto de Decreto de 1914 hablan de una enseñanza secundaria, impartida en las Universidades de provincia, pero con entidad propia, con unos objetivos específicos: bien la preparación para los estudios superiores, bien la obtención de conocimientos de cultura general. El plan de estudios de esta nueva segunda enseñanza contiene las dos áreas clásicas de conocimiento, Ciencias matemáticas, físicas y naturales y Literatura y artes, a la que se añade como innovación una tercera área, Ciencias morales y políticas, para cuyo estudio se crea, entre otras, una cátedra de Economía política y estadística. Así, después de mencionar sus múltiples y grandes aplicaciones el informe señala:

> *"...la Junta ha creído que debía terminar esta tabla de los estudios preparatorios de la juventud española por una cátedra en que bajo la dirección de un solo profesor se estudien los principios sistemáticos de las dos ciencias conocidas con el nombre de estadística y de economía política."*[1]

Aunque las Cortes no pudieron discutir este proyecto de Decreto, debido a la reacción y depuración absolutista a las que dio paso el golpe de Estado de Fernando VII en 1814 y que hicieron inviable la aplicación del mismo, es necesario resaltar la gran influencia de estos documentos en la legislación educativa posterior. Así durante el Trienio Constitucional (1820-1923) sirven de base al Reglamento General de Instrucción Pública de 29 de junio de 1821, Reglamento que establece en todas las Universidades de provincia la cátedra antes citada de Economía política y estadística. Sin embargo, el plan Duque de Rivas (1836) y el plan Pidal (1845) no contienen ninguna asignatura de Estadística.

En 1835, Tomás Serrano Server propone al Ministerio del Interior crear en Madrid una cátedra de Estadística, la clase de Comercio de la Sociedad Económica Matritense es requerida por dicho ministerio para emitir un informe sobre lo necesario y adecuado de aceptar tal petición. Después de oír la opinión poco favorable de la Comisión se decide no crear dicha cátedra, sino extender los estudios de Economía Política para que abarcaran las enseñanzas estadísticas necesarias.[2]

La creación de la primera cátedra de Estadística española, está ligada de alguna manera a la figura de Pascual Madoz e Ibáñez (1806-1870), conocido por su impresionante *Diccionario Geográfico-Estadístico-Histórico de España y sus posesiones de Ultramar*[3], obra de dieciséis tomos a la que dedicó siete años de su vida, de 1843 a 1850. Entre los cargos que ostentó están el de juez de Primera Instancia, gobernador del Valle de Arán, diputado por la provincia de Lérida, vicepresidente del Congreso de los Diputados y ministro de Hacienda. Madoz se relacionó con los más importantes estadísticos de su tiempo, como Quetelet y Moreau de Jonnes, de quien tradujo del francés la *Estadística de España.* En 1843 se le nombra presidente de una Comisión de Estadística, posición que aprovecha para proponer al Gobierno la creación de varias cátedras de Estadística, su proposición es rechazada, pero sirve de base a la Sociedad Económica de Amigos del País de Madrid para dar auge a esta disciplina. En palabras del profesor Sánchez-Lafuente:

---

[1] Clásicos Castellanos, B.A.E. Vol. XIX

[2] Eusebio Mª del Valle, miembro de esta comisión, llegaría a ser catedrático de Economía política y Estadística de la Universidad Central.

[3] MADOZ, P. (1845-1850): *Diccionario geográfico-estadístico-histórico de España y sus posesiones de Ultramar.* Estudio Literario-Tipográfico de P. Madoz y L. Sagasti, calle de la Madera Baja, 8. 16 tomos.

> *"La obra de Madoz consistió esencialmente en que creó una conciencia de la necesidad tanto del estudio teórico como de la ejecución de estadísticas por parte del Estado de una forma permanente. Su contacto con los estadísticos de la época no fue accidental, sino un intercambio de ideas sobre la forma de solucionar problemas de la realidad".*[4]

En el *Diccionario* de Madoz, bajo la palabra Madrid, se nos dice:

> "*Cátedra de Estadística (calle del Turco, número 9). La Sociedad Económica, a propuesta de varios individuos de su seno, estableció en 1844 esta útil enseñanza desconocida hasta entonces en nuestras escuelas: graves fueron las dificultades que hubo de vencer y, entre ellas, la de hallar persona idónea que por primera vez emprendiese esta clase de educación. Pero afortunadamente recayó la elección en nuestro recomendable amigo don José María Ibáñez, uno de sus socios y vocal secretario (a propuesta nuestra) de la Comisión de Estadística, creada por real decreto de 21 de agosto de 1843. El señor Ibáñez no pudo menos de acceder a las instancias de la sociedad y, encargado de esta cátedra, se vio precisado a dar a luz una obra elemental en la que prescindiendo de opiniones y sistemas en general, presenta los principios más esenciales y su aplicación a la práctica, indicando los diversos y multiplicados objetos a que deben dirigirse las investigaciones del estadista. Más aún, antes de la publicación de esta obra se instaló la cátedra, cuya apertura tuvo efecto en sesión pública celebrada el día 1º de diciembre del citado año y para cuyo curso fueron 38 los individuos matriculados que continuaron con eficaz asistencia y bastante aprovechamiento los dos cursos académicos que terminaron en 1846. Posteriormente, y con auxilio del libro de texto, se da un curso completo cada año, asistiendo en el presente (1848) 22 matriculados y un considerable número de oyentes. Es de desear, y nos consta, se tiene solicitado que el Gobierno dé a este estudio un carácter público, según los da a los de Economía política y Administración, parte integrante de estas ciencias."*[5]

Como acabamos de ver, la primera cátedra de Estadística española nace en el seno de la Sociedad Económica de Madrid en 1844, siendo su primer catedrático José María Ibáñez y Ramos. Poco después de establecida la cátedra el profesor Ibáñez escribe el libro titulado *Tratado elemental de Estadística, así en la parte filosófica y de teoría, como en la aplicación de sus principios a la práctica,* que, como indica en su portada, *está redactado con arreglo a las lecciones explicadas en la cátedra de dicha ciencia, establecida por la Sociedad Económica Matritense.* Este libro es el primer tratado de Estadística escrito por un español, ya que en esa época sólo existía en castellano la traducción, por Vicente Díez Canseco, del libro *Elementos de la Ciencia Estadística* del portugués Sampaio, editado en 1841 y traducido al castellano el mismo año y pequeños opúsculos sobre estudios estadísticos concretos.

Tras la muerte del profesor Ibáñez (7 de octubre de1856), el profesor Marcoartu pasa a encargarse de la cátedra de Estadística de la Sociedad Económica Matritense.

El 28 de agosto de 1850, siendo Ministro de Comercio, Instrucción y Obras públicas, Manuel de Seijas Lozano, se emite un Real Decreto reformando el plan de estudios vigente, por el que la Facultad de Filosofía queda dividida en las siguientes secciones: Literatura, Administración, Ciencias físico-matemáticas y Ciencias naturales. En la sección de Administración se imparte la asignatura de Estadística. Esta es la primera vez que en un plan de estudios se implanta la Estadística como asignatura independiente (aunque en la práctica se

---

[4] SÁNCHEZ-LAFUENTE FERNÁNDEZ, J. (1975): *Historia de la Estadística como ciencia en España (1500-1900).* I.N.E. Madrid, pág. 139.
[5] MADOZ, P. Tomo X, pág. 819.

sigue impartiendo junto con la Economía política y al principio se emplean los mismos libros de texto que para ésta), en la sección de Ciencias físico-matemáticas, sin embargo, no se estudia nada de Estadística.

El primer libro de texto de Estadística que se utiliza en las Facultades es la traducción al castellano del libro "*Elementos de Estadística*" de Moreau de Jonnes[6]

Las primeras tesis doctorales en Estadística

El 10 de septiembre de 1851, siendo Fermín Arteta Ministro de Comercio, Instrucción y Obras públicas, se emite una Real Orden con el reglamento para la ejecución del plan de estudios decretado el mes anterior.

Este Reglamento determina las nuevas pautas que han de seguirse para la obtención de los grados académicos.

En cuanto a la investidura del grado de doctor, los artículos 476 a 479 de dicho Reglamento nos dicen:

> *"El grado de doctor se conferirá siempre individualmente de la manera que sigue: el candidato escribirá una tesis sobre un punto cualquiera de la facultad o ciencia, y la imprimirá, entregando al Rector, con la anticipación de ocho días, el suficiente número de ejemplares para repartir al claustro. Llegado el día de la ceremonia, después de ser introducido en la sala por el padrino como en el caso de la licenciatura, leerá el impreso, que se distribuirá entre los circunstantes. Acto continuo le contestará uno de los catedráticos con un discurso relativo al objeto de la tesis y el modo con que la ha desempeñado, y en seguida el presidente le recibirá el juramento y le conferirá el grado con las insignias: hecho lo cual, se retirará acompañado del padrino y los bedeles después de abrazar a los doctores y de dar gracias al claustro...*
> *Los discursos y la tesis de que hablan los artículos anteriores se presentarán al Rector antes de leerse los primeros y de imprimirse la segunda, para que los revise y les ponga su visto bueno, sin cuyo requisito no se verificarán los actos".*

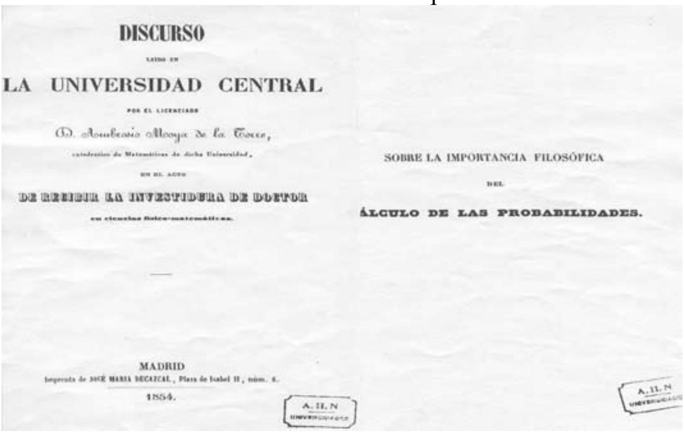
La tesis es, en este momento, un discurso que tenía que leerse en el acto de investidura como doctor, y no el resultado de un trabajo de investigación necesario para ser aprobado, como lo es en nuestros días. La primera vez que se menciona la palabra tesis en relación con el doctorado, es el 8 de julio de 1847, cuando, siendo Ministro de Comercio, Instrucción y Obras públicas, Nicomedes Pastor Díaz, se emite un real decreto en el que se modifica Plan Pidal y se incluye la lectura de una tesis como requisito indispensable a llevar a cabo en el acto de investidura de doctor.

Estando esta ley en vigor, D. Ambrosio Moya de la Torre, aprueba los ejercicios para la obtención del grado de doctor en la Facultad de Filosofía, sección de Ciencias Físico-Matemáticas, de la Universidad Central de Madrid.

El acto de la solemne investidura como doctor tiene lugar en la Universidad Central el 2 de julio de 1854, en dicho evento D. Ambrosio Moya de la Torre lee el discurso, previamente impreso para la ocasión por D. José Mª Ducazcal, titulado "*Sobre la importancia filosófica del cálculo de las probabilidades"* este discurso es considerado como la primera tesis

[6] La traducción es de IGNACIO ANDRÉS Y CASIMIRO PÍO GARBAYO DE BOFARULL (1857).

doctoral sobre Cálculo de Probabilidades leída en la Universidad Central de Madrid y por lo tanto la primera tesis doctoral sobre este tema leída en España.[7]


DISCURSO

LA UNIVERSIDAD CENTRAL

DE RECIBIR LA INVESTIDURA DE DOCTOR

SOBRE LA IMPORTANCIA FILOSÓFICA

ÁLCULO DE LAS PROBABILIDADES.

MADRID

1854.

Un año más tarde, el 15 de noviembre de 1855, D. Antonio Aguilar y Vela, catedrático de Matemáticas y Astronomía en la Universidad Central y secretario perpetuo de la Real Academia de Ciencias Exactas, Físicas y Naturales de Madrid, con motivo de su investidura como doctor, lee, en esa misma Facultad, la segunda tesis doctoral española sobre Cálculo de Probabilidades, titulada: "*De la importancia del estudio del Cálculo de Probabilidades"*.[8] Este documento es un buen indicador del estado en que se encontraba en aquella época el Cálculo de Probabilidades en nuestro país como señala uno de sus párrafos:

> *"Sensible es, Excmo. Señor, que en la Universidad Central, donde tan grande impulso ha recibido el cultivo de todas las ciencias, no se haya dado cabida al estudio del cálculo de las probabilidades, si no con la extensión que hoy abraza este importante ramo del análisis, al menos en sus teorías más elementales, en sus aplicaciones prácticas más comunes. Muchas son las ventajas que reportaría el país de la agregación de esta ciencia a alguna de las asignaturas de la sección de ciencias físico-matemáticas, y ya que de algunos años a esta parte se da tan justa importancia al estudio de las últimas, hágase lo mismo con sus principales aplicaciones, y entonces podrá mejor comprenderse la utilidad que en sí encierra el estudio de las ciencias exactas".*

La petición de D. Antonio Aguilar y Vela, al igual que la de otros matemáticos de la época, de que se estudie Cálculo de Probabilidades en alguna de las asignaturas de la sección de ciencias físico-matemáticas es desestimada. Como veremos a continuación, la nueva Ley Moyano, primera Ley de Instrucción pública española, que se establece en 1857 y que crea las Facultades de Ciencias Exactas, Físicas y Naturales, no incluye ninguna materia de Estadística ni de Cálculo de Probabilidades entre las asignaturas de estas Facultades, sin

[7] ESCRIBANO, M.C. Y BUSTO, A.I. (2004): "La primera tesis doctoral sobre Cálculo de probabilidades leída en la Universidad Central de Madrid", en *Historia de la Probabilidad y la Estadística II,* A.H.E.P.E., Delta, Madrid, págs. 287-299.

[8] BUSTO, A. I. Y ESCRIBANO, M .C. (2006): "D. Antonio Aguilar y Vela: sumisión del estudio del Cálculo de Probabilidades" en *Historia de la Probabilidad y la Estadística III,* A.H.E.P.E., Delta, Madrid, págs. 179-193.

embargo, sí se incluye a la Estadística como una de las asignaturas de la Facultad de Derecho en la sección de Administración.

Hay que esperar al 14 de marzo de 1933, para que en la Facultad de Ciencias de la Universidad Central se cree la asignatura Cálculo de probabilidades y Estadística matemática, primera cátedra de Estadística en una Facultad de Ciencias española.

La tercera tesis sobre Estadística leída en la Universidad Central es el discurso de D. Juan Cancio Mena e Irurzun, abogado de los Tribunales Nacionales y catedrático de la Escuela de Comercio de Santander, titulado *"Objeto e importancia de la Estadística"*[9], leído en la Facultad de Derecho en el acto de su investidura como doctor en Administración.[10]

En este discurso, D. Juan Cancio nos indica cómo para conocer el objeto de una ciencia es necesario primero definirla bien. Para darnos una buena definición se remonta al origen de la Estadística, que justifica en la necesidad de la administración, indicando que un buen gobierno debe conocer con exactitud todos los elementos de su país para poder establecer normas y disposiciones que le ayuden a prosperar. Así, nos dice:

> *"la Estadística es la ciencia que retrata física y moralmente la situación de un país, que define todos sus elementos, y que facilitando las vías de la prosperidad, abre ancho cauce al adelanto y al progreso."*

La tesis nos señala los datos indispensables que debe conocer un gobierno de su población: el número de los habitantes que componen los pueblos, sus alteraciones, el estado civil de sus individuos, la relación en la que se encuentran los sexos, las edades, los nacidos y los muertos. Para ello se debe elaborar un censo *"que registre minuciosamente la población y lleve cuenta detallada de todos sus movimientos"*, la formación del censo y su análisis detallado es obra de la Estadística. También ésta ayudará a conocer muchos otros elementos de un país que señalan la riqueza de un pueblo, como su topografía, sus tierras productivas y su clase de cultivo. D. Juan Cancio dice literalmente en su discurso:

> *"Por eso la administración inventariando la riqueza agrícola, dividiendo y clasificando los terrenos, y estudiando su naturaleza respectiva, puede evaluar exactamente la suma de todos los productos, sin recurrir a la hipotética base sobre la que levantaban sus cálculos algunos estadistas extranjeros, al tomar por tipo de sus operaciones el producto en bruto de una extensión determinada, al fijarse en el número de los instrumentos del cultivo o echar mano de otros sofísticos recursos para averiguar la suma total de la riqueza fisiocrática. Hé ahí las diferencias esenciales que existen entre la Estadística científica y la Estadística empírica, entre la que se remonta a la región sublime de los principios reclamando al auxilio de las demás ciencias, y la que se abandona a la práctica, entre la que funda sus investigaciones en sólidos argumentos y deduce consecuencias lógicas y naturales y entre la que siguiendo la ciega rutina se estaciona, se detiene, se paraliza, no da un paso por la vía del adelanto y de las mejoras. Hé ahí el carácter de la Estadística de los tiempos antiguos, de su estado primitivo, de su época naciente, hé ahí la fisonomía de la Estadística contemporánea, ilustrada, racional, filosófica".*

[9] Mena e Irurzun, J. C. (1860): *Objeto e importancia de la Estadística.* Discurso leído en la Universidad Central en el acto de investidura como doctor. Imprenta de Tomás Fortanet, Libertad 29, Madrid.
[10] Aunque Juan Cancio Mena licenciado en Filosofía, sección de Administración, cuando lee su tesis ya está en vigor la ley Moyano y los estudios de Administración se cursan en la Facultad de Derecho.

Nos dice también cómo la Estadística dará a conocer el verdadero estado de las industrias y cómo actuar para su adelanto.

A continuación expone la diferencia entre la Estadística y otras ciencias como la Economía política, la Aritmética política, la Geografía y la Historia. En cuanto a la segunda nos dice:

"La Aritmética política es una ciencia de hipótesis y a ellas confía sus investigaciones, pretende por ejemplo averiguar el número de habitantes de un país y recurre a los cálculos para resolver sus problemas, y en la mortandad, en el consumo o en otros fenómenos se fija para obtener sus resultados, mientras que la Estadística busca datos ciertos y seguros que alejen las dudas y presenten la verdad".

El profesor Mena termina su tesis con las siguientes palabras:

"El objeto de la Estadística es el conocimiento profundo de la sociedad apreciada en sus diversas fases de economía, situación y movimientos, y que es la ciencia de los hechos naturales, sociales y políticos expresados por términos numéricos".

Este discurso nos recuerda cómo la definición de la ciencia Estadística es uno de los objetivos que se plantean en el siglo XIX todos los autores de libros de Estadística.

La enseñanza de la Estadística a partir de la Ley Moyano

El 18 de marzo de 1857, siendo Ministro de Fomento Claudio Moyano Samaniego, se aprueba el Plan Orgánico de las Escuelas de Comercio y el Reglamento que ha de regir en las mismas. La enseñanza comercial se divide en dos periodos, en el primero, que dura tres años, se estudia una lección diaria de Geografía y Estadística industrial y comercial.

El 9 de septiembre del mismo año se promulga la primera Ley de Instrucción Pública, conocida como Ley Moyano. Esta ley crea las Facultades de Ciencias, en las que no se imparte Estadística, sin embargo, una de las asignaturas del cuarto curso de la Facultad de Derecho, necesaria para obtener el grado de Bachiller, común a todas sus secciones: Leyes, Cánones y Administración, es la de Economía y Estadística.

La Ley Moyano también regula las enseñanzas profesionales, entre las que se encuentra la de Profesor mercantil, una de cuyas asignaturas es Geografía y Estadística industrial y comercial.

Siendo Ministro de Fomento Rafael de Bustos y Castilla, un real Decreto del 26 de agosto de 1858 aprueba el Programa General de Estudios de Segunda Enseñanza, entre las asignaturas de aplicación a la agricultura, artes, industria y comercio se encuentra la Geografía y Estadística comercial, obligatoria para obtener el título de perito mercantil.

Pocos días después, el 11 de septiembre, siendo Corvera el Director General de Instrucción pública se reforma el plan de estudios de las Facultades.

La Facultad de Derecho se divide ahora en dos secciones, una de Derecho civil y canónico, y otra de Derecho administrativo

Para aspirar al grado de Bachiller en ambas secciones, se tiene que estudiar, entre muchas otras, la asignatura de Elementos de Economía política y de Estadística.

Este mismo día, se aprueba, mediante una Real Orden la lista de obras de texto para la segunda enseñanza, entre ellas se encuentra *"Geografía y Estadística industrial y comercial"* de D. Fabio de la Rada y Delgado (1858).

El día 25 del mismo mes se aprueban las listas de obras de texto para las Facultades. Los libros de texto que se debían utilizar en la Facultad de Derecho para impartir la asignatura Elementos de estadística son:

- *Tratado de estadística, por Mr. P. A. Dufau, traducido por Laroche y Sierra (1845)*
- *Elementos de estadística, por Alejandro Moreau de Jonnés (1857)*

El 26 de septiembre de 1861, se aprueban los libros de texto de segunda enseñanza. Las obras aprobadas para impartir *Estadística comercial* son:

- *Geografía comercial y estadística, por D. Gabino de Epalza (1861).*
- *Curso de Estadística elemental, por D. Fabio de la Rada y Delgado (1861).*

El 14 de septiembre de 1867 se aprueba como libro de texto para el estudio de la Estadística comercial en la Enseñanza Secundaria la obra *Curso de Geografía y Estadística industrial y mercantil, por D. Mariano Carreras y González (1863).*

Con la Revolución de septiembre de 1868, España entra en un periodo conocido como *El sexenio revolucionario (1868-1874)*. Se trata de un periodo muy inquieto en el que triunfan los principios del liberalismo radical y democrático, que se apoyan en los conceptos de soberanía nacional y sufragio universal.

Toda la legislación española, y por tanto la educativa, cambia de rumbo hacia el liberalismo. Es digno de notar que se liberaliza la enseñanza en todos sus grados y cualquiera que sea su clase, que todos los españoles quedan autorizados para fundar establecimientos de enseñanza, que los alumnos no tendrán que asistir a las clases para ser admitidos a examen, que para obtener los grados académicos no se necesitará estudiar un número determinado de años, sino las asignaturas que fijen las leyes, sufriendo un examen sobre cada una y el general correspondiente al grado. Los profesores podrán utilizar como libro de texto cualquiera que crean conveniente y no tendrán que presentar un programa de la asignatura. Se suprime la Facultad de Teología en las Universidades. Se suprime la investidura de los grados de Bachiller y de Licenciado. Los ejercicios del Doctorado podrán verificarse en todas las universidades y la investidura se hará sin exigir ningún juramento.

El 11 de febrero de 1873 se proclama la primera República, siendo Ministro de Fomento Eduardo Chao se aprueba un nuevo plan de estudios, conocido como el plan Chao.

Este plan, en un Decreto del 2 de junio, crea como Facultades independientes, en la Universidad de Madrid, las Facultades de Filosofía, de Letras, de Matemáticas, de Física y Química y de Historia Natural, esperando poder ampliarlas al resto de Universidades.

Sin embargo, pese a todos los cambios citados, el estudio de la Estadística permanece como estaba en el anterior periodo político.

Tras el sexenio revolucionario y la llegada de la Primera República, en España comienza una nueva etapa conocida como la Restauración, es decir, la vuelta al trono de la dinastía borbónica.

Aunque nada más comenzar este periodo se da un paso atrás en los principios de libertad de enseñanza, al amparo de la Constitución de 1876 y su principio de libertad de creación de centros docentes, un grupo de profesores liberales funda la Institución Libre de Enseñanza, una especie de universidad libre desde la que se impulsa la renovación pedagógica y cultural de la España de la época.

La nueva Constitución obliga al Gobierno a retocar la normativa vigente en materia de educación, por eso, el 13 de agosto de 1880, siendo Ministro de Fomento, Fermín de Lasala y Collado se emite un Real Decreto introduciendo varias reformas en el plan de estudios, entre las novedades se encuentra la introducción de ejercicios prácticos en las diferentes carreras, sobre todo en las de Ciencias, Medicina y Farmacia, y el dedicar más atención a los trabajos de laboratorio, tan indispensables en algunas disciplinas. Es un primer intento para iniciar a los estudiantes en el difícil mundo de la investigación.

También van en este sentido algunas de las reformas que se realizan en las diferentes Facultades desde 1883 (Derecho) hasta 1886 (Farmacia). En alguna éstas se exige a partir de ahora un nuevo requisito para la obtención del grado de doctor: la lectura de una tesis sobre un punto doctrinal o de investigación que ha de ser calificada por un tribunal y defendida por el doctorando. A partir de ahora la tesis deja de ser un mero discurso leído en el acto de investidura como doctor para tímidamente empezar a ser un trabajo serio de investigación original.

Bibliografía

AGUILAR Y VELA, A. (1855): "*De la importancia del estudio del Cálculo de las Probabilidades"*. Discurso leído en la Universidad Central. Imprenta de ANCOS. Madrid.

BUSTO CABALLERO, A. I.; ESCRIBANO RÓDENAS, M. C. (2002): "Primeros intentos para la organización de la enseñanza de la Estadística en España: cursos de estadística y sus aplicaciones 1.950-1952" en *Historia de la Probabilidad y de la Estadística*. A.H.E.P.E. Madrid. Págs. 193 – 204.

BUSTO, A. I. Y ESCRIBANO, M .C. (2006): "D. Antonio Aguilar y Vela: sumisión del estudio del Cálculo de Probabilidades" en *Historia de la Probabilidad y la Estadística III,* A.H.E.P.E., Delta, Madrid, págs. 179-193.

BUSQUETA, J .J.; PEMÁN, J. (2002): *"Les universitats de la Corona d´Aragó, ahjir i avui".* Estudis històrics. Pòrtic. Biblioteca Universitaria. Universitat de Lleida.

Colección Legislativa De España. Imprenta Nacional. Madrid.

CRUZ MUNDET, J. R. (Ed.) (2003): "*Archivos universitarios e historia de las universidades"*. Universidad Carlos III. Editorial Dykinson. Madrid.

ESCRIBANO RÓDENAS, M. C.; BUSTO CABALLERO, A. I. (2002): "La creación en España de la primera Escuela de Estadística" en *Historia de la Probabilidad y de la Estadística*. A.H.E.P.E. Madrid. Págs. 205 – 220.

ESCRIBANO RÓDENAS, M.C.; BUSTO CABALLERO, A.I. (2004): "La Primera Tesis Doctoral sobre Cálculo de Probabilidades leída en la Universidad Central de Madrid", en *Historia de la Probabilidad y la Estadística II*. A.H.E.P.E. Madrid. Págs. 287-300.

ESPAÑOL GONZÁLEZ, L. (2004): "La primera oposición a Cátedra de Estadística Matemática en la universidad española", en *Historia de la Probabilidad y la Estadística II,* A.H.E.P.E., Delta. Madrid. Págs. 387-400.

GARMA PONS, S. (1.990): "Las Matemáticas en España en la primera mitad del siglo XX" en *Actas de las XV Jornadas Luso-Espanholas*. Vol. VI. Didáctica e Historia da Matemática. Universidad de Evora. Págs. 3 – 65.

GARMA PONS, S.; SÁNCHEZ RON, J. M. (1.989): "La Universidad de Madrid y el Consejo Superior de Investigaciones Científicas" en *Revista Alfoz*. Págs. 59-77.

GIL DE ZÁRATE, A. (1855). *De la instrucción pública en España*. Impresión del Colegio de Sordomudos. 3. Varios Tomos. Madrid.

MADOZ, P. (1845-1850): *Diccionario geográfico-estadístico-histórico de España y sus posesiones de Ultramar.* Estudio Literario-Tipográfico de P. Madoz y L. Sagasti, calle de la Madera Baja, 8. 16 tomos.

MARTÍN PLIEGO, F. J. (1.997): "Notas sobre la historia de la probabilidad en España" en *Revista Zubía* , nº 15. Logroño. Págs. 155-167

MARTÍN PLIEGO, F .J. (1.997): "Historia de la probabilidad en España" en *Revista de Historia Económica.* Año XV nº 1. Universidad Autónoma de Madrid. Págs. 161-184.

MENA E IRURZUN, J. C. (1860): *Objeto e importancia de la Estadística.* Discurso leído en la Universidad Central en el acto de investidura como doctor. Imprenta de Tomás Fortanet, Libertad 29, Madrid.

MOYA DE LA TORRE, A. (1854): "*Sobre la importancia filosófica del cálculo de las probabilidades"*. Discurso leído en la Universidad Central. Imprenta de José María Ducazcal. Madrid.

MIGUEL ALONSO, A. (2003): "Los estudios de doctorado y el inicio de la tesis doctoral en España 1847-1900", en *Archivos universitarios e historia de las universidades.* Universidad Carlos III. Editorial Dykinson. Madrid. Págs. 197-217.

NEGRÍN FAJARDO, O. (2005): *Veinticinco ensayos de historia de la educación española moderna y contemporánea.* UNED. Madrid.

VEA MINUESA, F. (1.995): *Las matemáticas en la enseñanza secundaria en España en e siglo XIX.* Cuadernos de Historia de la Ciencia. Universidad de Zaragoza.

VIÑAO FRAGO, A. (1.982): *Política y educación en los orígenes de la España contemporánea. Examen especial de sus relaciones en la enseñanza secundaria*. Editorial siglo XXI. Madrid.

Capítulo 8

El Instituto Nacional de Estadística de Lisboa

Un recorrido a través de su historia: orígenes, creación, funcionamiento y evolución

GABRIELA FERNÁNDEZ BARBERIS
Mª DEL CARMEN ESCRIBANO RÓDENAS
Universidad CEU San Pablo

Organismos dispersos, trabalhando sob a inspiração de princípios diferentes, por vezes colhendo elementos nos mesmos campos e para fins muitas vezes próximos, mas seguindo métodos diversos e desconhecendo-se mutuamente, não podiam corresponder ao que deles a estatística devia exigir e ao que, pelo seu custo, a colectividade podia esperar.
(Armindo Monteiro, O *Instituto Nacional de Estatística,* Lisboa, 1936)

Introducción

La Estadística es un bien público que, a lo largo de la historia, ha sido demandado cada vez más por los distintos segmentos de la sociedad. Del poder público a la sociedad civil, a los agentes económicos y al ciudadano, todos procuran conocer mejor y más rápidamente la realidad económica, social y ambiental en la que están inmersos, en la que actúan y de la que son parte integrantes, es decir, demandan la información estadística que la caracteriza. La mayor exigencia por parte de los múltiples usuarios se junta con el creciente engrandecimiento de fenómenos, hechos y realidades sobre los cuales se exige producir información estadística creíble, sistemática e imparcial.

Portugal, que actualmente forma parte de la Unión Europea, ha ido adecuándose a la sociedad contemporánea y ha ido modernizando su sistema estadístico con el paso del tiempo, hasta llegar al actual Instituto Nacional de Estadística, que si bien es el único en todo el país, se lo denomina en múltiples ocasiones como el INE de Lisboa.

Reseña histórica de la institucionalización de la Estadística en Portugal

En la evolución de las estadísticas oficiales portuguesas existen algunas fechas claves que deben tenerse en cuenta para comprender el Sistema Estadístico Nacional de Portugal, conocer a fondo su historia y valorar la gran labor desempeñada por los historiadores.

El primer censo de población conocido de lo que actualmente es Portugal (aproximadamente la región entre los ríos Duero y Guadiana, entonces parte de la provincia Romana conocida como Lusitania) fue realizado en el año del nacimiento de Cristo, el año cero de nuestra era, por orden del emperador César Augusto quién mandó realizar un censo de todas las provincias romanas.[1]

Durante el período medieval, desde el siglo XII al siglo XIV, en los primeros días en que Portugal se constituyó como nación, la información que podría considerarse cuantitativa o estadística era fragmentaria, irregular y muchas veces expuesta a interpretaciones ambiguas y contradictorias. Hubo algunos intentos de realización de censos de población en los años 1260, 1279, 1421/22 y 1527.

Tras un largo período de letargo, es en 1636 cuando comienzan a utilizarse algunos métodos para cuantificar la realidad social, no obstante siguen careciendo de sistematicidad y precisión. En el período comprendido entre los años 1636 – 1774 deben destacarse: *The Review of the People of War (1636)* y *The List of Houses and Souls within the Lands of Portugal (1732)*, también conocido como el Censo de Población del Marqués de Abrantes.

Desde finales del siglo XVIII hasta mediados del siglo XIX, las estadísticas fueron fortaleciéndose e intensificándose cada vez más, hicieron su aparición en el campo comercial y se expandieron hacia demografía y otras áreas de la vida social y económica (el año 1775 marcó el comienzo de la producción anual de la publicación *Balanzas Generales del Comercio del Reino de Portugal y de sus Dominios,* 1798 el del *Censo de Población de Pina Manique,* y 1801 el del *Censo de Población del Conde de Lindares*).

La primera institución estadística oficial que existió en Portugal fue la "Superintendência Geral dos Contrabandos e Descaminhos dos Reais Direitos nestes Reinos e seus Domínios" (siendo Contador General, Mauricio José Teixeira de Morais) que elaboró a partir de 1775, en el campo de las estadísticas del Comercio, las "Balanças Gerais do Comércio do Reino de Portugal com os seus Domínios e com as Naço es Estrangeiras"[2].

Así pues, y sin olvidar la creación en 1815 de la "Comissa*ã*o de Estatística e Cadastro do Reino", es preciso llegar a 1841, con el consiguiente atraso en comparación con el contexto europeo[3], para que el gobierno atribuyera a la Estadística un valor estratégico, y considerara la

[1] Así consta la referencia en la Biblia (Lucas 2).

[2] Publicación editada anualmente, en forma manuscrita, entre 1796-1831, en 3 volúmenes: uno sobre comercio con los Dominios, otro con las Naciones Extranjeras, y otro, resumido, con los Dominios y las Naciones Extranjeras. El INE de Portugal posee una colección en la que sólo faltan los volúmenes de 1798, 1808 y 1822, que, están repartidos, respectivamente, en la Biblioteca Municipal de Santarém, en el Archivo Histórico del Ministerio de Obras Públicas y en la Biblioteca de Río de Janeiro.

[3] Tuvieron el primer organismo oficial de Estadística: Suecia 1749, Francia 1800, Alemania 1805, Holanda 1826, Austria 1829. En España, el inicio de la estadística oficial tuvo lugar en 1856 con la creación de la Comisión General de Estadística del Reino, aunque el primer intento de crear este organismo fue la denominada "Oficina de Balanza" fundada en 1802, con intentos previos de D. Pedro López de Lerena en 1789. Véase: Fernández Barberis, G.; Escribano Ródenas, M.C. (2004): *Participación española en las primeras reuniones internacionales de estadística;* en: Historia de la Estadística y la Probabilidad (II), A.H.E.P.E., pp. 420-416, Delta Universidad. Madrid.

necesidad de creación de un organismo para que desempeñara una actividad de producción y difusión de información estadística oficial de un modo *sistemático y suficientemente amplio.*

El Ministro del Reino Rodrigo da Fonseca Magalhães somete a aprobación el Decreto de 30 de Abril de 1841, por el cual se crea en la "Inspecçã o-Geral de Obras Públicas la Secçã o de Estatística e Topográfica" con el objetivo de:

> *"Establecer um método permanente e seguido de alcançar, reunir, e coordenar os dados positivos, informaçõ es e esclarecimentos que constituem a Estatística do Reino, tã o indispensável para servirem de base a muitas das providências administrativas, assim como aos projectos, empresas, e trabalhos tendentes a promover a prosperidade pública",* que todavía, *"face à escassez de recursos postos à disposiçã o, nunca pôde ultrapassar o fosso entre as ambiçõ es do seu mandato e as possibilidades de sua realizaçã o".*

El año 1841 puede considerarse como el de la creación del primer servicio oficial de estadística, coincidiendo, curiosamente, con el año en que tuvo lugar el comienzo de la enseñanza de la Estadística a nivel universitario en Portugal, en la Facultad de Derecho de la Universidad de Coimbra. El Prof. Adriã o Pereira Forjaz de Sampaio, fue quien introdujo la enseñanza de la Estadística en Portugal con su obra *"Primeiros Elementos da Ciência Estatística"*, que a su vez utilizó para sus clases suplementarias al Curso de Economía Política que también impartía en la Facultad de referencia.

Es digno de destacar que esta obra del Prof. Adriã o de Sampaio fue la base del *primer libro de Estadística publicado en España* a través de una traducción del Prof. Vicente Diez Canseco, cuyo prólogo data del 27 de Agosto de 1841. A este hecho se refiere el Prof. Juan Sánchez-Lafuente Fernández en su *"Historia de la Estadística como Ciencia en España de 1800 a 1900"* afirmando que: *"En aquel tiempo sólamente existía en español un tratado de Estadística , el del portugués Sampaio".*

En 1857, por el Decreto de 8 de Agosto, fue creada en el Ministerio de Obras Públicas una "Comissã o Central de Estatística do Reino" con el propósito de "organizar o plano geral da estatística em todos os ramos da administraçã o pública, superintendendo na sua execuçã o e centralizando a publicaçã o dos trabalhos estatísticos", transfiriéndose así la actividad estadística oficial desde el Ministerio del Reino hacia el Ministerio de Obras Públicas.

La Ley de 6 de Junio de 1859 dio autorización al gobierno para reformar el Ministerio de Obras Públicas, Comercio e Industria, con la obligación expresa de crear una "Repartiçã o de Estatística", cuya materialización se reflejó en el Decreto de 5 de Octubre del mismo año, creándose en la Dirección de Comercio e Industria la "3ª Repartiçã o de Estatística".

El Decreto de 28 de Diciembre de 1864 creó el "Conselho Geral de Estatística do Reino" (constituido de facto el 5 de Abril de 1866), presidido por el Ministro de Obras Públicas. Entre otras, se le atribuyeron funciones de: discusión y adopción de reglas aplicables a los métodos de recogida de datos estadísticos; análisis, estudio y comparación de resultados obtenidos en las diversas investigaciones estadísticas; suministro de información a los ministerios. El Reglamento del citado Consejo se publicó el 24 de Abril del mismo año, cuya importancia fundamental consistió en establecer su división en once Secciones Permanentes, de las cuales quedaban fuera las estadísticas del Comercio Exterior y las Estadísticas Agrícolas, éstas totalmente inexistentes y las primeras confiadas a los servicios aduaneros.

El Decreto de 23 de Julio de 1864 (del Duque de Loulé) ordenó la realización en Portugal del primer "Recenseamento da Populaçao" con las características técnicas de general, nominal y simultáneo, que llevó a cargo la Repartición de Estadística, pero cuya ejecución debió enfrentarse a numerosas dificultades de carácter financiero debidas a una notoria insuficiencia de recursos.

El Decreto de 31 de Diciembre de 1868 reformó una vez más el Ministerio de Obras Públicas, Comercio e Industria, procediéndose a la reorganización de la Repartición de Estadística de la Dirección General de Comercio e Industria que se dividió en dos Secciones.

Al año siguiente, el Decreto de 16 de Diciembre, extinguió el Consejo General de Estadística del Reino y en su lugar se repuso la anterior Comisión Central de Estadística, eliminándose las dos Secciones en que estaba dividida, reduciéndose el número de sus miembros y definiéndose de manera más precisa las funciones que le competían en su calidad de organismo de dirección y orientación superior de la actividad estadística oficial; dotándola de gran capacidad operativa.

La reforma iniciada en 1869 se completó en 1885 a través de la creación de las "Comissões Distritais de Estatística" (Decreto de 19 de noviembre), con sede en las capitales de distrito y compuestas por las siguientes entidades y personas: gobernador civil; presidente de la comisión ejecutiva de la Junta General; un concejal de la Cámara Municipal de la capital de distrito; un profesor de liceo; el agrónomo del distrito; el intendente de pecuaria; y un ciudadano indicado por el gobernador civil, cuyo funcionamiento dependía de las instrucciones de la Comisión Central de Estadística.

El Decreto de 28 de Julio de 1886 alteró la estructura orgánica de la Repartición de Estadística de la Dirección General del Comercio e Industria (creada en 1859), dividiéndola en tres secciones, y se otorgó al gobierno autoridad para reorganizar la Comisión Central de Estadística.

En 1887, por el Decreto de 3 de Febrero, la Comisión Central de Estadística fue transformada en "Conselho Superior de Estatística", ampliándose el número de competencias respecto de las cuales podía emitir su parecer. Por este mismo Decreto se estableció, por primera vez en Portugal, el *principio de obligatoriedad de respuesta* a todas las encuestas estadísticas realizadas por la Repartición de Estadística (Principio de Autoridad Estadística). Sin embargo, no fueron tipificadas las correspondientes transgresiones ni tampoco fueron estipuladas las sanciones a las eventuales transgresiones[4].

En el Ministerio de Obras Públicas, el 4 de Junio de 1892, se nombró una comisión para elaborar un proyecto general que organizase los servicios oficiales de estadística del país, cuyo trabajo vino a culminar en una relación con base en la cual, por el Decreto Nº 5 de 1 de Diciembre, fue reorganizada la Estadística adoptando un *criterio descentralizador.*

En el año 1898, mediante el Decreto de 30 de Junio, se produjeron dos alteraciones de importancia. Por un lado, la creación de la "Direcçao-Geral de Estatística e dos Próprios Nacionais", y por otro, el traslado de la tutela de la Estadística desde el Ministerio de Obras Públicas hacia el Ministerio de Hacienda, reconociéndose así *"a necessidade de centralizar organismos que até aí tinham andado dispersos e a sua má colocaçao dentro do quadro geral dos serviços públicos".*

[4] A pesar de esto, dicho principio fue suprimido en 1929 por el Decreto Nº 16.943 de 7 de Junio y cuando habían transcurrido nada menos que 42 años desde su institucionalización.

Durante el año 1911 esta Dirección General sufrió distintos cambios de denominación hasta que, por último, pasó a llamarse solamente, "Direcçāo-Geral de Estatística". Esta última fue reorganizada por el Decreto Dictatorial Nº 5.524 de 8 de Mayo, que reorganizó las Direcciones Generales del Ministerio de Finanzas en general. Pero, aún así, la eficacia de esta reforma no se alcanzó hasta el año siguiente, por el Decreto Nº 6.607 de 10 de Mayo de 1920, que organizó y reglamentó la Dirección General de Estadística, donde toda la organización prevista se fundamentaba en la supervisión del Consejo Superior de Estadística, destinado a orientar superiormente la actividad estadística nacional. El citado Consejo sólo se reunió en dos ocasiones, 1923 y 1926, y fue en ésta última en la que se discutió una propuesta de Reforma Estadística que fue presentada por el propio Director General de Estadística[5], pero que no tuvo aceptación posterior. En 1929 se opera una nueva Reforma Estadística, a través de los títulos siguientes:

1. Decreto Nº 16.369 de 15 de Enero: creó el billete estadístico aduanero y reformó las viejas formas de recogida estadística del comercio exterior;
2. Decreto Nº 16.537 de 23 de Febrero: renovó la producción de las estadísticas demográficas y extinguió la Inspección de Demografía y Estadística de la Dirección General de Salud, pasando sus atribuciones a la Dirección General de Estadística;
3. Decreto Nº 16.538 de 23 de Febrero: creó el servicio de las publicaciones y notas de la Dirección General de Estadística;
4. Decreto Nº 16.927 de 1 de Junio: organizó la estadística de las transmisiones de propiedad inmobiliaria e hipotecas y de la de las sociedades comerciales;
5. Portaria Nº 6.288 de 1 de Junio: organizó la estadística de los documentos comerciales y la de las quiebras y los concordatos;
6. Decreto Nº 16.943 de 7 de Junio: dio eficacia al principio de autoridad estadística, estableciendo las transgresiones estadísticas y las respectivas penas, así como también organizando el proceso a seguir para su punición.

Evolución de los Principios Orientadores del Sistema Estadístico Nacional

El Instituto Nacional de Estadística (INE) de Lisboa

El Instituto Nacional de Estadística (INE) de Lisboa, se creó bajo la tutela del Ministro de Finanzas, y estableció por primera vez, de un modo sistemático, *los Principios Orientadores* del Sistema Estadístico Nacional. Inicialmente, se le atribuyeron funciones de anotación, elaboración, publicación y comparación de los elementos estadísticos referentes a los aspectos de la vida portuguesa que interesaran a la Nación, al Estado o a la Ciencia.

El autor del proyecto de la citada ley fue el Prof. Armando Rodrigues de Sttau Monteiro[6] que en ese momento desempeñaba las funciones de Director General de Estadística, en cuyo cargo había permanecido de 1928 a 1935, y en 1939 desempeñó las funciones de Director General del recién creado Instituto Nacional de Estadística.

Cabe destacar que esta reforma de 1935 no se ocupó de la estadística oficial de la Colonias y hubo que esperar a la creación del INE para que se contemplase la *Estadística Colonial.*

---

[5] Vitorino Henriques Godinho, Teniente Coronel del Cuerpo del Estado Mayor y Profesor de la Escuela Militar, Ministro de Negocios Extranjeros en 1924 y Ministro del Interior en 1925, fue Director General de Estadística de 1922 a 1928.

[6] Profesor Catedrático de la Facultad de Derecho de Lisboa. Tuvo una vida política muy activa: Subsecretario de Estado de Finanzas; Ministro de las Colonias; Ministro de Asuntos Exteriores; Embajador de Portugal en Londres.

El INE reestructuró los Servicios Oficiales de Estadística con vista a dotarlos de los medios necesarios a una acción más eficiente, y definió por primera vez, de forma sistemática, los *Principios Básicos Orientadores del Sistema Estadístico Nacional*, principios esos que fueron los siguientes:

- Centralización de los Servicios.
- Autonomía Técnica.
- Autoridad Estadística y Responsabilidad de los Informadores.
- Secreto Profesional.
- Atribución al INE del derecho de fiscalización de las publicaciones estadísticas producidas por otras entidades.

A pesar de la auténtica revolución que supuso el INE en el panorama estadístico nacional portugués, la verdad es que los resultados obtenidos, a medida que los años pasaron, dejaron de corresponder a las expectativas y a las nuevas exigencias subsiguientes de las políticas de desarrollo económico, principalmente, a partir de 1953, con el Primer Plan de Fomento.

Con el paso del tiempo la eficacia y el funcionamiento del INE se deterioraron, sobre todo por razones de naturaleza política.[7]

En la década de los sesenta, cuando se da inicio a uno de los períodos de mayor desarrollo económico del Portugal Contemporáneo, la Ley Nº 2123 de 14 de Diciembre de 1964 sobre la organización y ejecución del Plan Intercalar reconoce, la necesidad de proceder a una inmediata y profunda reestructuración del Sistema Nacional de Estadística, indispensable al planeamiento de "todo el espacio portugués". Esta reorganización tendrá que esperar aún un par de años.

La Reforma de 1966

El Decreto-Ley Nº 46.925, de 29 de Marzo de 1966, promulgó la reorganización del Sistema, el cual fue complementado por el Decreto Nº 46.926 de la misma fecha, aprobanó el R*egulamiento del Sistema*, cuyos proyectos se deben al Dr. Amaro Duarte Guerreiro, entonces Director General del INE, cargo que ejerció de 1955 a 1973[8] .

Con esta nueva ley se procuraba dar un nuevo aliento a los Servicios de Estadística, de acuerdo con los objetivos siguientes:

- Codificar y actualizar la legislación vigente, alguna con más de 30 años de existencia;
- Reafirmar los principios de la ley 1911, principalmente el de la centralización estadística;
- Introducir el principio de coordinación estadística, creando a tal efecto el Consejo Nacional de Estadística, con amplia representación de los servicios públicos y de los sectores privados (se recuperaba, de esta forma, la tradición del Consejo Superior de Estadística, que, en 1935, se rechazara);
- Organizar a escala nacional el sistema estadístico, para tornar eficiente la orientación técnica y la coordinación por el Consejo Nacional y por el Instituto Nacional de Estadística;

---

[7] Por ejemplo, en el Censo de 1940 aparecieron tasas de analfabetismo que se retuvieron para no desacreditar las Conmemoraciones de ese año y en la década de los cincuenta, la policía política detuvo a algunos especialistas del INE y penetró varias veces en sus instalaciones para revisar los despachos y secretarías de los funcionarios.
[8] En 1955-1964 como Director General Sustituto, en 1964-1965 como Director Interino y en 1965-1973 como Director General.

- Dotar los servicios con medios humanos y materiales, asegurando su formación y perfeccionamiento técnico.

El Sistema Estadístico Nacional gozaba así de una estructura más fluida, compuesta por: Consejo Nacional de Estadística, Comisiones Consultivas de Estadística, INE y Órganos Delegados del Instituto.

El INE continuaba gozando, por ley, de completa autonomía técnica, adjudicándose exclusivamente el ejercicio de las funciones de anotación, purificación, coordinación y publicación de datos estadísticos, responsabilizándose por las funciones de ejecución de inquisiciones y todos los otros trabajos estadísticos, pero omitía el de la planificación.

El Consejo Nacional de Estadística expresó repetidamente la necesidad de revisión de los medios y modos de acción del Instituto, hasta que, en Diciembre de 1972, fue aprobada por el Consejo la relación de una Comisión implicada en el estudio de las líneas maestras informadoras de la organización y funcionamiento del INE.

La Reforma de 1973

La reestructuración efectuada en 1973, por el Decreto-Ley Nº 427/73 y el Decreto Nº 428/73, ambos de 25 de Agosto, no afectó a los principios básicos del Sistema Estadístico Nacional, por considerarse que eran los más adecuados al caso portugués. Las alteraciones se verificaron, fundamentalmente, a nivel de la organización de los servicios del INE, y fueron, en síntesis, las siguientes:

- Refuerzo de la capacidad de dirección del Instituto;
- Reestructuración del sector de los estudios estadísticos;
- Creación de nuevas direcciones de servicios y su dotación;
- Resolución de problemas de personal relevantes – gestión, especialización de funciones, organización de cursos, movilidad del personal, etc.
- Creación de delegaciones regionales del Instituto.

Después de la resolución de 25 de Abril de 1973, instalada la democracia pluripartidaria, abandonado definitivamente el imperio colonial, superada la inestabilidad política que se hizo sentir durante el período revolucionario y ultimadas las negociaciones para la adhesión de Portugal a la Comunidad Económica Europea, consumada en 1985, se tornó imperiosa una efectiva reforma estructural del Sistema Estadístico Nacional, dada su importancia para la toma de decisiones del más diverso orden.

En junio de 1986, el gobierno resolvió crear una Comisión de Reestructuración del Sistema Estadístico Nacional, con el mandato de proceder a un diagnóstico riguroso de la situación y preparar, adoptando un horizonte temporal de cinco años, un conjunto de acciones teniendo en vista una mayor operatividad del sistema.

De acuerdo con la relación efectuada por la Comisión de Reestructuración, presidida por Manuel Vilares, el Sistema Estadístico Nacional enfermaba, fundamentalmente, de los siguientes estrangulamientos:

- un inadecuado reglamento jurídico, excesivamente rígido;
- la propia composición del Consejo Nacional de Estadística, limitada casi exclusivamente a miembros de la Administración Pública;
- excesiva rigidez en las normas de secreto estadístico;

- abandono de las funciones de coordinación;
- el Estatuto del INE, organismo público sencillo, sin autonomía;
- la escasez de cuadros superiores especializados en el dominio de la concepción, tratamiento y análisis de la información estadística.

La Reforma de 1989

En base a las conclusiones de esta relación efectuada por la Comisión de Reestructuración, a través de la Ley 6/89, de 15 de Abril y del Decreto-Ley 280/89, de 23 de Agosto, se procedió a la reforma del Sistema Estadístico Nacional que se operó, esencialmente, en cuatro dominios:1) el del secreto estadístico; 2) el de la centralización; 3) el de la coordinación y, 4) el de la autoridad estadística.

A nivel orgánico, el Sistema Estadístico Nacional pasa a comprender tan sólo dos órganos, el Consejo Superior de Estadística que substituyó al Consejo Nacional de Estadística, reforzando las atribuciones y competencias, y ampliando su composición a representantes más allá de la Administración Pública, principalmente representantes de las confederaciones patronales y sindicales, de la Asociación de Municipios, del Banco de Portugal, de las asociaciones de consumidores, y del *Instituto Nacional de Estadística*, ampliamente reformulado y con incumbencia para celebrar protocolos con instituciones universitarias con vista a la creación de cursos para formación de cuadros superiores de estadística. Este último pasa a ser ahora "un instituto público dotado de personalidad jurídica, autonomía administrativa y financiera y patrimonio propio, teniendo por objeto funciones de anotación, averiguación, coordinación y difusión de datos estadísticos que interesen al País".

Se procuró una alteración radical en la filosofía de gestión del Instituto de Estadística, de forma que la componente económica y la componente financiera pasasen a intervenir clara y directamente en las decisiones; incentivar la producción estadística en la perspectiva de los usuarios, conferir movilidad a los medios y flexibilidad al funcionamiento interno y relaciones al exterior, posibilitando la adecuación de la gestión a las características del proceso de obtención de productos estadísticos y similares a un proceso industrial típico; y reforzar la capacidad institucional. El cuadro orgánico del INE pasó a estar estructurado de la siguiente forma:

- Dirección, con un presidente y otros dos elementos nombrados por decisión conjunta del primer miembro y del "Ministro do Planeamento e Ordenamento do Território", a quien compete dirigir la actividad en general y representar al INE;
- Consejo de Administración, compuesto por la Dirección y cuatro miembros más sin funciones ejecutivas, de nombramiento gubernamental, con competencia para definir y acompañar la orientación general y las políticas de gestión del INE;
- Comisión de Fiscalización, con tres elementos nombrados por el Ministro de Finanzas para examinar periódicamente la situación financiera del INE, verificar la ejecución de las decisiones del Consejo de Administración y emitir los juicios convenientes en este dominio.

Conclusiones

El INE surgió en 1935, en el contexto de modelo centralizado del Sistema Estadístico Nacional, en una tentativa de dar respuesta a la demanda, cada vez mayor, de información estadística, no sólo por parte de la Administración Pública sino también de las entidades particulares.

Desde esta fecha, se han sucedido diferentes reformas en el Sistema Estadístico Oficial de Portugal, entre las que cabe resaltar las de 1966, 1973 y 1989.

Sólo esta última, puede considerarse como un hito en la historia del Sistema Estadístico Nacional, se consiguió introducir una verdadera reforma de fondo, que concedió al INE personalidad jurídica, autonomía administrativa y financiera y patrimonio propio.

El desafío del INE portugués es afirmarse como un centro de excelencia a escala nacional e internacional. Por tanto, deberá contar con el lugar único y específico que ocupa en el seno de las instituciones nacionales y con el empeño, profesionalismo y dedicación de todos los que en él desempeñan su actividad profesional y sus capacidades intelectuales y humanas.

Bibliografía

CARVALHO MACHADO, J. (1985): *I.N.E. 50 Años. Portugal 1935-1985.* Imprensa Nacional. Casa da Moeda, E.P.

CELESTINO REY, F. (2002): "La Génesis Administrativa del INE, 1939-1948". En: *Historia de la Probabilidad y la Estadística (I),* pp. 181-191. A.H.E.P.E. Editorial AC. Madrid.

DE SOUSA, F. (1995): *História da Estadística em Portugal.* INE, Litografía Amorum, Lisboa.

FERNÁNDEZ BARBERIS, G.; ESCRIBANO RÓDENAS, M.C. (2004): "Participación española en las primeras reuniones internacionales de Estadística". En: *Historia de la Probabilidad y la Estadística (II),* pp. 404-416. A.H.E.P.E. Delta Universidad. Madrid.

FERREIRA DA CUNHA, A.S. (1995): *O Sistema Estatístico Nacional. Algumas notas sobre a evoluçao dos seus princípios orientadores: de 1935 ao presente.* Instituto Nacional de Estadística, Portugal.

FERREIRA DA CUNHA, A.S. (2000): *The Portuguese National Statistical System. A Brief Overview.* Instituto Nacional de Estadística, Portugal.

FERREIRA DA CUNHA, A.S. (2001): *Nótulas Históricas em Torno do Sistema Estatístico Nacional.* Instituto Nacional de Estadística, Portugal.

IMPRENSA NACIONAL DE LISBOA (1936): *O Instituto Nacional de Estadística*.

INSTITUTO NACIONAL DE ESTADÍSTICA, Portugal (2006): *Setenta anos 1935-2005. O Instituto Nacional de Estatística ao Serviço da Sociedade Portuguesa.* INE, Portugal.

SANCHEZ LA FUENTE, J. (1975): *Historia de la Estadística como Ciencia en España (1500 – 1900).* Tesis Doctoral, Universidad de Málaga. INE.

Capítulo 9

Tablas y curvas de mortalidad en el siglo XVIII

Semejanzas y divergencias*

MARC BARBUT
Centre d'Analyse et de Mathématiques Sociales, EHESS, París

La comparación de tablas mortalidad no es ninguna novedad. Antoine Deparcieux lo hizo ya en su « Ensayo» de 1746, a propósito de unas tablas bien conocidas en su época: Petty, Halley, Kerseboom, John Smart (esta última conocida también como Tabla de Hogdson).

Su intención era construir una tabla mejor que todas ellas, que sirviera para una población ideal y fuera aplicable de la manera más amplia posible a lo que hoy llamamos el cálculo actuarial.

Años más tarde, E. Beauvisage (1867), A. Quicquet (1893) – [13] y [15] de la bibliografía – y otros autores compararon tablas entre sí con el mismo objetivo que Deparcieux.

Aquí, sin embargo, mi objetivo es otro. No tengo nada que demostrar, ni pretendo aproximarme a un ideal; sino que, de forma más modesta, quiero señalar algunos hechos. En cuanto a las teorías, cada uno de nosotros es lo suficientemente inteligente para extraer de los hechos lo que le conviene para pensar.

Pese a que no soy demógrafo ni historiador, me he arriesgado a construir dos tablas empíricas, las dos últimas del Cuadro I.

Enseguida veremos el porqué y la justificación de este enfoque algo insólito.

* Traducido por María Gómez y José MªArribas Dep. de Sociología I, UNED, Madrid

1- Cómo leer una tabla de mortalidad

Comencemos por examinar la tabla [6] creada por Lambert, en § 24, que figura aquí en el Anexo I. Nos referiremos a ella a partir de ahora como Lambert 24 (cf. nota 4, Figura I). Tiene 7 columnas y es, que yo sepa, la más completa de todas las tablas del siglo XVIII.

Leemos en la primera columna la edad x expresada en años, en la segunda el número m(x) de fallecidos entre la edad x y la edad x+1, y en la tercera, el número y(x) de supervivientes de edad x.

Dado que toda persona viva de edad x morirá en ese mismo año, o en cualquier año ulterior, tenemos así las relaciones:

(1.1) $m(x) = y(x) - y(x + 1).$

(1.2) $y(x) = \sum_{i=0}^{\infty} m(x + i).$

N.B. Aquí (véase columna 3), $y(x + i) = 0$ cuando $x + i \geq 103$.

CUADRO I

Lista de las tablas estudiadas

Autores	Fecha de publicación	Localización de las observaciones	Tipo de población y número de observaciones	Período de las observaciones
Halley (1)	1693	Breslau (Silesia)	Conjunto de la población (en torno a 15.000)	1678-1691
Deparcieux	1746	Toda Francia	*Tontiniers** y órdenes religiosas	1689-1742
Hogdson *J. P.* (2) *Lambert* (3) theórico	1747 1759 1772	Londres Londres ___	Conjunto de la población (en torno a 270.000) (en torno a 750.000)	1728-1737 1728-1757 ___
Lambert 24 (4)	1772	Alemania (?)	?	Antes de 1770
Duvillard	1796 1806	Toda Francia	Conjunto de la población (en torno a 101.000)	Antes de 1789
Fourier	1821	París	Censo (en torno a 700.000)	1817
Beauvisage (Caisse Lafarge)	1867	Toda Francia	*Tontiniers* (en torno a 39.000)	1793-1864
Dictionnaire *Robert* (5) Abbé de *Choisy* Mémoires	1977 1727	Europa, Francia Francia	Personalidades célebres (en torno a 1.350) Aristocracia (en torno a 250)	1667-1801 1630-1740

Notas sobre el cuadro:

* Tontiniers eran los miembros de una mutua aseguradora, la tontine, en la que cada miembro aportaba una suma destinada a convertirse en una renta vitalicia al llegar a una cierta edad. La persona que sobrevivía a esa edad recibía las sumas aportadas por aquellos que habían fallecido.

1. Salvo en el caso de la tabla de Fourier (cf. [10] de la bibliografía), se muestra para cada año de observación, la edad de defunción de aquellos que mueren ese año concreto.
2. La « Tabula X » de Süssmilch, reproducida en el Anexo III *infra*, es un resumen de esta tabla que figura *in extenso* en la *Collection of yearly bills*... (cf. [4]).
3. Véase Anexo II *infra*. Resulta de la modelización de la « Tabula X ».
4. Reproducida en el Anexo I *infra*. No sabemos a partir de qué datos la ha elaborado Lambert (cf. [6] y [7]).
5. Diccionario « Le petit Robert des noms propres ».

Los números de la 3ª columna se calculan sumando los de la 2ª empezando por el último (a tergo, como decía B. Pascal).

En cuanto a la 4ª columna, los números Y(x) que la constituyen, son las sumas (siempre comenzando desde el último), de las y(x + i) :

$$(1.3) \qquad Y(x) = \sum_{i=0}^{\infty} y(x + i).$$

Luego Y(x) es la población total de edad superior o igual a x. Concretamente, Y(0) es la población total sobre una base de 10.000 nacimientos. En este caso:

$$Y(0) = 295.022.$$

Esto significa que para conocer la población total de la tabla de mortalidad Lambert 24 basta con conocer el número anual de nacimientos y multiplicarlo por 29,5: éste es el multiplicador que fue tan útil en el siglo XVIII para estimar la población de diversos territorios.

Pero 29,5 es también la esperanza de vida al nacer, como indica la columna 6.

De forma más general, los números de la sexta columna indican para cada edad x la vida media (denominada también por abuso del lenguaje, esperanza de vida v(x) a esa edad). Lambert la calcula por medio de los de las columnas 3 y 4.

En efecto, por definición, tenemos (de forma aproximada, ya que todos los muertos de edad x no fallecen el mismo día):

$$v(x) = \frac{\sum_{i=0}^{\infty}(x+i+1)\, m(x+i)}{\sum_{i=0}^{\infty} m(x+i)}.$$

Sea, ateniéndose a (1.1) y (1.2) supra:

$$(1.4) \qquad v(x) = \frac{\sum_{i=0}^{\infty}(x+i+1)[y(x+i)-y(x+i+1)]}{y(x)}.$$

Tras la simplificación y considerando (1.3), el numerador de (1.4) es:

$$x\,y(x) + \sum_{i=0}^{\infty} y(x+i) = x\,y(x) + Y(x),$$

finalmente:

(1.5) $$v(x) = x + \frac{Y(x)}{y(x)}.$$

Por ejemplo, para la edad de 30 años, tenemos según la tabla Lambert 24 :

$$v(30) = 30 + \frac{127.090}{4.517} = 30 + 28{,}136 \simeq 58{,}1.$$

N.B. La mayor parte de los autores (Hogdson, Deparcieux, …) eligen como edad de muerte la media $(x + \frac{1}{2})$, lo que da una mejor aproximación de v(x). De esta manera obtenemos:

(1.5)′ $$v(x) = x + \frac{Y(x+1)}{y(x)} + \frac{1}{2}.$$

El número:

(1.6) $$s(x) = v(x) - x = \frac{1}{2} + \frac{Y(x+1)}{y(x)},$$

se denomina supervivencia media a la edad x. (Para Lambert, (v(x) – x) es igual a $\frac{Y(x)}{y(x)}$; por lo tanto, la esperanza media es según su cálculo: $s(x) + \frac{1}{2}$).

En cuanto a la columna 7, indica la vida mediana (también denominada vida probable) p(x) a la edad x, definida por:

(1.7) $$y[p(x)] = \frac{1}{2}\,y(x).$$

Por ejemplo, para x = 30, se obtiene según la tabla:

$$y[p(30)] = \frac{4.517}{2} = 2.258{,}5.$$

Este último número está comprendido entre 2.271 (58 años) y 2.175 (59 años) ; luego p(30) es por lo tanto ligeramente superior a 58 años, 58,1 como se puede ver claramente en la columna 7 para x = 30.

Queda por explicar la quinta columna: representa las inversas de las tasas de mortalidad $\tau(x)$ para cada edad x, a saber:

$$\tau(x) = \frac{m(x)}{y(x)} = \frac{y(x) - y(x+1)}{y(x)}.$$

Las obtenemos consecuentemente dividiendo los números de la columna 2 por aquellos de la columna 3. Así, para 30 años tenemos:

$$\tau(x) = \frac{72}{4.517} = \frac{1}{62{,}736} \simeq \frac{1}{63} \simeq 1{,}6\%\ .$$

Igualmente, a 63 años:

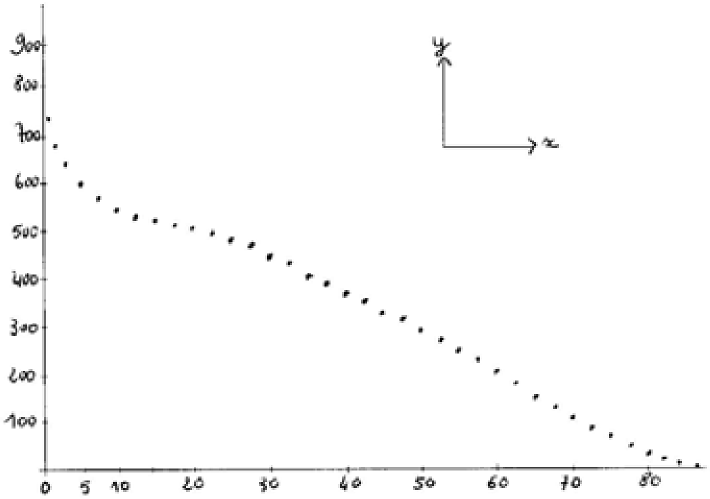
$$\tau(x) = \frac{105}{1.773} = \frac{1}{16,886} \simeq \frac{1}{17} \simeq 5,9\% .$$

N.B. Todos estos cálculos tienen sentido únicamente suponiendo que la población es estacionaria, es decir, que las y(x) no varían en el transcurso del tiempo.
Veremos (§ 2) algunas indicaciones relativas al caso en que el número de supervivientes de edad x en fecha t sea una función y(x,t) de las dos variables x y t.

Dos observaciones complementarias al leer la tabla Lambert 24 :

1) La mortalidad m(x) – columna 2– tiene el mínimo 33, para 17 años y el máximo 105, para 63 años. La curva de mortalidad (cf. Figura 1) tiene por lo tanto dos puntos de inflexión correspondientes para x = 17 y x = 63.

FIGURA 1. **Curva de mortalidad de la tabla *Lambert 24*.**
x = edad en años, $y(x)$ = número de vivos de edad x.



2) De 25 años hasta aproximadamente 55 años, la mortalidad varía poco, en una « horquilla » de 70 a 90 muertos por año. La curva de mortalidad será por lo tanto prácticamente una línea recta entre esas dos edades. Esto ya lo había intuido Leibniz; vamos a retomarlo.

N.B. Algunas tablas sólo constan de dos columnas, la de la edad x y la del número de vivos y(x) (columnas 1 y 3 de la tabla Lambert 24). De hecho, todas las otras columnas pueden calcularse a partir de esas dos.

2- Cómo construir una tabla (empírica) de mortalidad

Como se ha señalado más arriba, el número de vivos de edad x, en una población dada, varía por lo general con el tiempo t. Es una función y(x, t) de la edad x y de la fecha t.

Sean:

y(x, t): el número de vivos de edad x el 1 de enero del año t.

m(x, t): el número de muertos de edad x a lo largo del año t.

Así tenemos :

(2.1) $$m(x, t) = y(x, t) - y(x + 1, t + 1).$$

Como todos los vivos de edad x en la fecha t tienen que morir necesariamente a esa edad o a una edad ulterior, resulta obvio que:

(2.2) $$y(x, t) = m(x, t) + m(x + 1, t + 1) + \ldots = \sum_{i=0}^{\infty} m(x + i, t + i).$$

(2.2) es además una consecuencia inmediata de la ecuación (2.1).

(2.2) explica por qué a la función y(x, t), que, siendo t constante, es la pirámide de edades en fecha t, se la denomina también tabla (o curva) de mortalidad.

Teniendo ahora dos fechas t_0 y t_1, con:

$$t_0 < t_1 \quad \text{y} \quad t_1 - t_0 = T.$$

1) A través de los registros civiles, o de cualquier otra fuente, revelamos, para cada edad x, el número de muertos observados entre esas dos fechas. Sea:

(2.3) $$M_{01}(x) = m(x, t_0) + m(x, t_0 + 1) + \ldots + m(x, t_1 - 1) = \sum_{i=0}^{T-1} m(x, t_0 + i).$$

2) Aplicando la ecuación (2.2) de arriba, se calcula el número total de muertos de edad ≥ x durante el período observado. Sea:

$$Y_{01}(x) = M_{01}(x) + M_{01}(x + 1) + \ldots = \sum_{j=0}^{\infty} M_{01}(x + j),$$

es decir:

(2.4) $$Y_{01}(x) = \sum_{j=0}^{\infty} \sum_{i=0}^{T-1} m(x + j, t + i) = \sum_{i=0}^{T-1} \sum_{j=0}^{\infty} m(x + j, t + i) = \sum_{i=0}^{T-1} \sum_{j=0}^{\infty} (y(x + j, t + i) - y(x + j + 1, t + i + 1)).$$

Tras la simplificación, obtenemos finalmente:

(2.5) $$Y_{01}(x) = \sum_{i=0}^{T-1} y(x, t_0 + i) - \left[\sum_{j=1}^{\infty} y(x + j, t_1) - \sum_{j=1}^{\infty} y(x + j, t_0)\right].$$

En (2.5), el primer término es la suma del número de vivos de edad x para todo el período observado.

En cuanto al término complementario entre corchetes, es la variación del número de vivos de edad > x entre la fecha inicial y la fecha final del período observado. Teniendo en cuenta las relaciones (1.5), (1.6) del párrafo precedente, este término puede expresarse también mediante el número medio de supervivientes a través de:

$$\mathrm{y}(\mathrm{x},\ t_1)\ \mathrm{s}(\mathrm{x},\ t_1) - \mathrm{y}(\mathrm{x},\ t_0)\ \mathrm{s}(\mathrm{x},\ t_0).$$

Si la pirámide de edades de la población estudiada ha variado poco a lo largo del período, tendremos así de forma aproximada:

$$(2.6) \qquad Y_{01}(x) \simeq \sum_{i=0}^{T-1} y(x, t_0 + i) = T\, \bar{y}_{01}(x),$$

donde $\bar{y}_{01}(x)$ es el número medio de vivos de edad x observado para todo el período estudiado.

En el caso de una población estacionaria (§ 1), (2.6) sería teóricamente una igualdad. Pero la estacionaridad no es más que una hipótesis, que, como todas las hipótesis, puede ser verificada sólo de manera aproximada por los datos empíricos.

La determinación de una tabla empírica de mortalidad se obtiene eligiendo una edad de base – por ejemplo, 0 la edad al nacer – y normalizando la tabla. Planteamos por lo tanto:

$$(2.7) \qquad \frac{Y_{01}(x)}{Y_{01}(0)} \times A = y(x),$$

donde A es un coeficiente de escala, generalmente 1.000 ou 10.000, menos frecuentemente 100.000: y(x) es por lo tanto el número observado de supervivientes de edad x para A nacimientos.

3- Aplicación a la construcción de dos « micro-tablas »

Con objeto de establecer comparaciones en relación a la mortalidad entre unas poblaciones ordinarias – es decir, que reúnan todas las categorías sociales – y unas poblaciones social o geográficamente diferenciadas, es necesario disponer de tablas construidas por medio de observaciones realizadas en tal categoría determinada.

Por lo que respecta a las categorías de privilegiados, la tabla de Deparcieux, construida sobre una población de tontiniers y miembros de órdenes religiosas fallecidos entre finales del siglo XVII y mediados del siglo XVIII (cf. Cuadro 1, supra), responde en parte a esta exigencia. Esto es aún más cierto en el caso de la tabla de Beauvisage, constituida únicamente por tontiniers, pero ésta concierne principalmente al siglo XIX.

Luego he construido dos tablas con mis propios medios, que son escasos.

3.1- La Tabla « ROBERT DE LOS NOMBRES PROPIOS»

Del diccionario « le petit Robert des noms propres » (ediciones « Le Robert », París, 1977), he anotado la edad de fallecimiento de todas las personalidades europeas que figuran en este diccionario (familias soberanas, papas, grandes militares, hombres políticos, etc … pero también escritores, artistas, …) muertas entre 1667 y 1801. De esta forma, he obtenido una tabla construida a partir de cerca de 1300 observaciones si contamos sólo las personalidades fallecidas de « muerte natural »; esta tabla figura en el Anexo IV.

Si le añadimos las muertes violentas, especialmente numerosas al final del período (Revolución francesa), tenemos entonces casi 1450 nombres.

Digo « casi» porque seguramente he cometido algunos errores en este fastidioso trabajo de copista.

He dispuesto las columnas de la tabla Robert del Anexo IV de la misma manera que las de la tabla Lambert 24 del Anexo I, pero omitiendo la columna que concierne a las tasas de mortalidad: no tendría ningún sentido teniendo en cuenta lo pequeños que son los números en juego.

Por la misma razón, podemos reemplazar la tabla del Anexo IV por una tabla con datos ajustados de manera que se atenúe el efecto de las fluctuaciones aleatorias. Lo he hecho depurando los datos brutos a través de una media móvil de 11 puntos. Salvo para edades muy altas, dicha media no varía mucho, como lo demuestra el Cuadro II aquí debajo, extraido de la tabla depurada. La diferencia relativa entre estas dos no supera nunca el 3%.

CUADRO II

Edad	30	40	50	60	70	80	90
Datos brutos	1278	1222	1111	910	547	209	25
Datos depurados	1277,4	1224,8	1108,9	901,8	559,3	212	30,7

Obviamente, la comparación entre esta tabla y las otras tendrá sentido sólo a partir de una edad distinta de la de nacimiento: hace falta haber vivido bastante tiempo para poder llegar a ser célebre. Nosotros tomaremos aquí como base la edad de 40 años.

3.2- Estacionaridad de la *Población Robert*

Es remarcable que esta limitada población (aproximadamente unos 1300 individuos), observada en un largo período (135 años), presente al menos tres características que permitan considerarla estable.

Primera característica: la estructura de la población es prácticamente la misma al principio y al final del período (cf. § 2 y ecuación 2.5 supra).

Efectivamente, tenemos para el período inicial (indexado 0) 1667 – 1673, y para el período final (indexado 1) 1798 – 1801, 74 y 67 observaciones respectivamente. Distribuimos las observaciones por tramos de edad (al morir) quinquenales.

Obtenemos, siguiendo las mismas notaciones que para la tabla del Anexo IV, el Cuadro III abajo.

No hay ningún fallecido menor de 25 años, ni mayor de 94 años. Las columnas 6, 7 y 8 muestran que las dos curvas de mortalidad normalizadas se acercan mucho una a la otra: la diferencia no supera nunca el 5%, de lo que, según un test como el de Kolmogorov-Smirnov, se deduce que en un umbral del 5%, así como en un umbral del 1 % hay una misma « distribución genérica ». En cuanto a las poblaciones agregadas (columnas 9 y 10) son prácticamente idénticas. De lo que se deduce que en la ecuación (2.5) el término complementario entre corchetes es insignificante.

CUADRO III

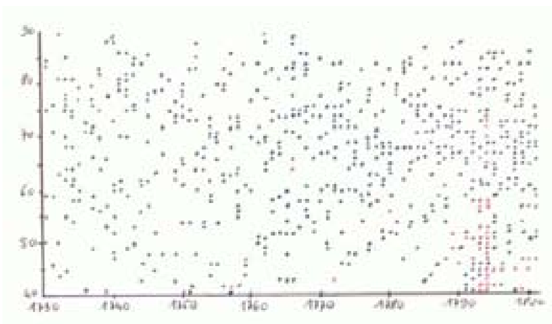
X	m_0	m_1	y_0	y_1	y_0 %	y_1 %	y_1% - y_0%	Y_0 %	Y_1 %	x
25-29	0	3	74	67	100	100	0	100	100	25-29
30-34	2	0	74	64	100	95,5	- 4,5	88	87,8	30-34
35-39	3	3	72	64	97,3	95,5	- 1,8	76,1	76,25	35-39
40-44	2	2	69	61	93,25	91	- 2,25	64,5	64,6	40-44
45-49	6	4	67	59	90,55	88	- 2,55	53,4	53,55	45-49
50-54	5	5	61	55	82,45	82,1	- 0,35	42,6	42,85	50-54
55-59	7	3	56	50	75,7	74,6	- 1,1	32,7	32,85	55-59
60-64	8	7	49	47	66,2	70,15	4	23,7	23,8	60-64
65-69	14	16	41	40	55,5	59,7	4,2	15,8	15,2	65-69
70-74	13	12	27	24	36,5	35,8	- 0,7	9,2	8	70-74
75-79	3	5	14	12	18,9	17,9	- 1	4,8	3,6	75-79
80-84	7	6	11	7	14,85	10,45	- 4,4	2,6	1,45	80-84
85-89	3	1	4	1	5,4	1,5	- 3,9	0,8	0,2	85-89
90-94	1	0	1	0	1,35	0	- 1,35	0,15	0	90-94
1	2	3	4	5	6	7	8	9	10	

N.B. Al contrario, la comparación entre el primer período quinquenal (1728-1732) y el segundo (1753-1757) de la tabula X de Süssmilch (cf. Anexo III) muestra, para la ciudad de Londres, una débil pero significativa diferencia en un intervalo de tan sólo 30 años.

Segunda característica: no hay correlación entre la edad x al morir y la fecha t de la misma. Esto es patente en la Figura 2 a continuación.

FIGURA 2. **Diagrama de (no)-correlación entre la fecha de fallecimiento (en abscisas) y la edad al morir (en eje de ordenadas) 1730 à 1801.**

(En rojo: los fallecidos por muerte violenta-ver el CD).



<u>Tercera característica</u>: las dos características principales de la distribución de la edad de los fallecidos (media, desviación) son estables cuando desglosamos las observaciones en 9 períodos de 15 años. Es lo que muestra el Cuadro IV.

CUADRO IV

Diccionario *Robert des noms propres.* Fallecimientos entre 1667 y 1801.

(Características de 9 períodos de 15 años)

Período	*n*	*v*	max	min	σ	*r*	*a*	*n'*	*v'*
1667-1681	149	64,55	96	20	14	0,15	0,5	155	63,9
1682-1696	152	63,9	90	21	16,5	0,13	0,5	154	63,75
1697-1711	120	66,1	94	31	13,5	0,03	0,1	120	66,1
1712-1726	126	67	95	24	17	- 0,07	- 0,03	130	66,3
1727-1741	122	66,3	93	15	13,7	0,03	0,1	124	66,1
1742-1756	107	67,2	93	29	16,6	- 0,1	- 0,4	109	66,9
1757-1771	136	65,45	100	18	14,3	- 0,05	- 0,04	145	64
1772-1786	158	65	87	29	13,8	0,08	0,2	163	64,4
1787-1801	225	64	96	25	14,4	- 0,1	- 0,3	337	57,3
Total	1295	65,3	100	15				1437	63,22

n : número de observaciones, muertes naturales

n' : número de observaciones, incluyendo las muertes violentas

v : edad media al morir (por muerte natural)

v' : edad media al morir (incluyendo las muertes violentas)

σ : desviación

r : coeficiente de correlación linear entre *x* y *t*

a : coeficiente de regresión de *x* en relación a *t*

N.B. La aparente estabilidad de la población Robert sugiere que para los « privilegiados», y en esta época, no hay a lo largo del período ni avance ni retroceso de la mortalidad. Por otra parte, una de las razones por las que he construido las tablas Robert y Choisy ha sido comprobar esta hipótesis. ¿Cómo sería en poblaciones mucho más numerosas de « privilegiados » ? La cuestión me parece abierta.

3.3- La tabla CHOISY

El abate de Choisy (1644-1724) publicó unas Memorias para servir a la historia de Luis XIV (publicadas en 1727 y reeditadas en Mercure de France, colección Temps retrouvé 7, París, 1965). En ellas cita cerca de 245 personalidades –no garantizó el número exacto – pertenecientes todos a la alta aristocracia y muertas (todas por causas confusas) entre 1623 y 1749. A través de las reseñas biográficas redactadas por el editor, sabemos la edad del fallecimiento, y podemos elaborar con ella una tabla, demoninada Choisy en lo sucesivo, análoga a la tabla Robert.

Esta tabla – en datos brutos – se haya en el Anexo V. Aquí el número de observaciones es demasiado pequeño para que podamos realizar los mismos tests de estacionaridad ya empleados para los 1.300 individuos de la tabla Lambert.

No obstante, el gráfico « fecha de fallecimiento – edad al morir», que no reproduzco aquí, presenta, aunque con menor densidad en la nube de puntos, los mismos rasgos de dispersión y de ausencia de tendencia que el de la Figura 2 más arriba.

Por el contrario, el análogo del Cuadro IV más arriba, donde se dividió casi todo el período en dos partes (grosso modo, los siglos XVII y XVIII) parece mostrar una ligera tendencia al aumento de la longevidad en el siglo XVII (veasé Cuadro V), y un estrechamiento de la dispersión en el siglo XVIII. Teniendo en cuenta lo pequeños que son los números, sería arriesgado sacar conclusiones de este hecho.

CUADRO V

Período	n	v	max	min	σ	r	a
1630-1699	124	63,9	91	20	13	0,5	0,3
1700-1749	108	68,2	97	27	8,2	0,02	- 1,15
Total	232	65,9					

En cambio, es muy significativo que esta micro-tabla de mortalidad se ajuste casi igual de bien al modelo de Gompertz y con unos valores del mismo orden de magnitud para los parámetros (aquí anticipo lo que veremos en § 6) que las otras tablas que vamos a estudiar, aunque éstas hayan sido construidas con decenas, incluso con centenares, de miles de observaciones.

Es lo que muestran las columnas 7, 8 y 9 del Cuadro Anexo V.

Se puede observar el mismo fenómeno, incluso aún más acentuado, en la tabla Robert. En este caso, creo que está perfectamente justificada la inclusión de las dos micro-tablas en la lista de las que se estudian aquí.

4- Diferencias, semejanzas y correlaciones. Estudio empírico

Estamos ahora en condiciones de hacer algunas comparaciones, situándonos en primer lugar exclusivamente desde el punto de vista de los datos empíricos ; más adelante (§ 5 et 6), lo haremos sobre la base de ajustes de las tablas empíricas a unas leyes teóricas.

Algunas tablas de la lista del Cuadro I, p. 2 – la tabla de J. P. (alias « Süssmilch – Tabula X »), la tabla Lambert-teórica que procede directamente de ella, y la misteriosa Lambert 24 – se reproducen aquí en los anexos III, II y I respectivamente.

Las tablas de Halley, Deparcieux y Hogdson se encuentran principalmente en [8], p. 276, 278 et 280 respectivamente. Las de Duvillard, Beauvisage y, de nuevo, Deparcieux en [13], p. 18 y ss.; finalmente, la de Joseph Fourier (censo de la población de París de 1817) en [10] (Cuadros anexos relativos a la población, cuadro 8º).

N.B. Al comparar los datos de esta última tabla con los otros debemos tener en cuenta siempre que no se trata de una tabla de mortalidad construida en torno a todo un período, sino de los resultados de un censo realizado en un momento dado (en este caso, el año 1817). Por lo tanto, no podría compararse directamente con las otras tablas salvo en un régimen estable; y al leerla, la palabra vida en las expresiones « vida media », « supervivencia media », etc... debe ser siempre reemplazada por la palabra edad.

4.1- Juventud y vejez. Dos grupos y un caso aparte

El Cuadro VI presentado aquí debajo destaca claramente la singularidad de la tabla Fourier. Las tablas de Halley, Deparcieux, Beauvisage, Robert y Choisy no figuran en este cuadro, porque no comienzan más que a las edades de 1, 3, 3, 15 y 20 años respectivamente.

CUADRO VI

	Número de vivos por 1.000 al nacer				
Edad	*Lambert 24*	*Duvillard*	*Hogdson*	*J. P*	*Fourier*
3	644	625	564	597	--
5	604	583	526	550	932
10	554	551	490	516	866,6
20	512,5	502	459	485	692,5
40	375	335	294	312	342
	Tablas « universales »		Londres		Censo
	¿Alemania ?	Francia	1728-1737	1728-1757	París 1817

Comentarios.
- La semejanza entre las tablas Hogdson y J. P. y el hecho de que en la segunda los jóvenes sean más numerosos que en la primera, no debe sorprender: se trata de la misma gran ciudad, durante dos períodos de los cuales el segundo incluye al primero, y éste es posterior para los dos últimos tercios.
- Las tablas « universales » Lambert 24 et Duvillard son también cercanas una a la otra, y tienen más jóvenes que las de Londres. Es normal: estas tablas no incluyen sólo a los ciudadanos, y probablemente los datos fueron recopilados más tarde en el siglo XVIII. Por otra parte, las dos tienen además en común que nadie sabe exactamente con qué datos han sido construidas.
- En 1817, la ciudad de París tiene muchos más jóvenes que lo que indicarían las tablas construidas en el siglo precedente. Pero hacia los 40 años las curvas confluyen, y más allá de esta edad, París tiene muchos menos viejos.

El Cuadro VII de supervivencias medias para las mismas edades permite desglosar de forma nítida dos grupos:

- un grupo de « privilegiados »; en este grupo, el hecho de que Beauvisage supere a Deparcieux se explica fácilmente. El período de las observaciones es muy posterior, y la tabla Beauvisage contiene sólo datos de tontiniers, mientras que Deparcieux incluye un tercio de tontiniers y dos tercios de miembros de órdenes religiosas;
- un grupo de « el común de los mortales», en el que las dos tablas londinenses son superadas por las otras tres, incluida la de Halley, aunque ésta concierne a una población urbana y en un período muy anterior; sin duda alguna, en Breslau las condiciones de vida eran mucho mejores que en Londres.

Destacamos por último el descenso de las edades medias en París-1817, ya desde la edad de 5 años.

CUADRO VII

Supervivencia media en diversas edades *x*. Juventud

Tabla \ x	0	3	5	10	20	40	Período
Choisy	--	--	--	--	--	29,75	1623-1749
Robert	--	--	--	--	--	28	1667-1801
Beauvisage	--	54,65	53,6	49,8	42,4	28,45	1793-1864
Deparcieux	--	47,65	48,25	46,8	40,25	27,5	1689-1742
Duvillard	28,8	42,5	43,4	40,8	34,25	22,9	Avant 1789
Lambert 24	29,5	41,5	42,2	39,9	33,8	22,3	Avant 1770
Halley	--	39,7	41,2	40,4	34,2	22,2	1678-1691
Hogdson	23,85	38,5	39,3	37	29,2	19,55	1728-1737
J. P.	25,6	38,35	39,5	37	29	19,25	1728-1757
Fourier	32,7	--	29,9	27	22,4	15,5	1817

Después de la juventud, la vejez.

Los análogos Cuadros VII’ y VI’ muestran las mismas diferencias entre los privilegiados y « el común de los mortales».

CUADRO VII’

Supervivencia media en diversas edades. Vejez

x	*Choisy*	*Beauvisage*	*Robert*	*Deparcieux*	*Duvillard*	*Lambert 24*	*Halley*	*Hogdson*	*J. P.*	*Fourier*
40	29,75	28,5	28	27,5	22,9	22,3	22,1	19,5	19,25	15,5
60	13,75	14	13,5	14,25	12	11,8	12,4	12,5	12,2	8,6
80	5,5	5	4,8	4,7	4,6	5,7	4,5	5,7	5,3	ε

Se observa que si las diferencias entre los dos grupos son aún nítidas a los 60 años, desaparecen a los 80 años; y que, incluso a esa edad, el censo de París de 1817 se distingue por su singularidad: en particular, ε significa «casi cero».

CUADRO VI’

	Número de vivos por 1.000 a 40 años					
x	50	60	70	80	Población	Fechas
Choisy	935	814,5	539	228,5	Alta aristocracia	1623-1749
Beauvisage	902,5	756,8	516	203	*Tontiniers*	1793-1864
Robert	909	744,5	447,5	171	Personalidades	1667-1801
Deparcieux	800,5	705	472	179,5	*Tontiniers* y miembros de órdenes religiosas	1689-1742
Duvillard	805,5	578,5	292,5	94,75	Universal – Francia	Avant 1789
Lambert 24	793	555	291,5	91	Universal - ?	Avant 1770
Halley	778	544,5	319,5	91,5	Breslau	1678-1691
Hogdson	694	442	235	97	Londres	1728-1737
J. P.	689	433	231	86,5	Londres	1728-1757
Fourier	619	308	96	14,3	París	1817

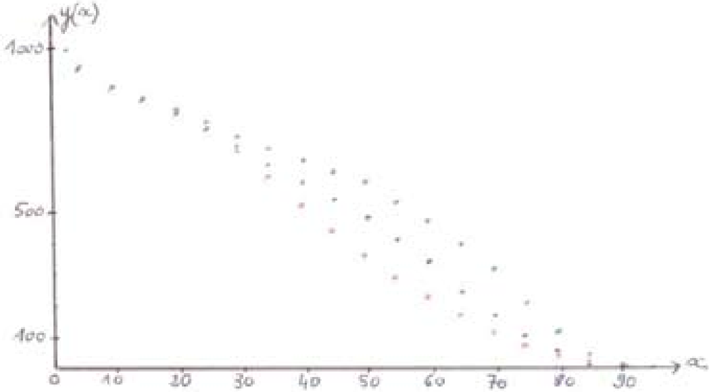
Advertencia. Comparado con el Cuadro VII’, este cuadro hace patente que si todas las tablas (excepto la Fourier) dan aproximadamente 5 años de esperanza de vida a los 80 años, los privilegiados que sobreviven a esa edad son dos veces más numerosos que el resto.
De nuevo, la población de París en 1817 se distingue por el muy bajo número de vivos mayores de 80 años, e incluso de 70 años.

4.2- Semejanzas y diferencias en algunos gráficos

Examinemos ahora las tres curvas de mortalidad de la Figura 3. Se refieren a tres poblaciones muy diferentes tanto por el período de observación como por la categoría de sujetos observados.

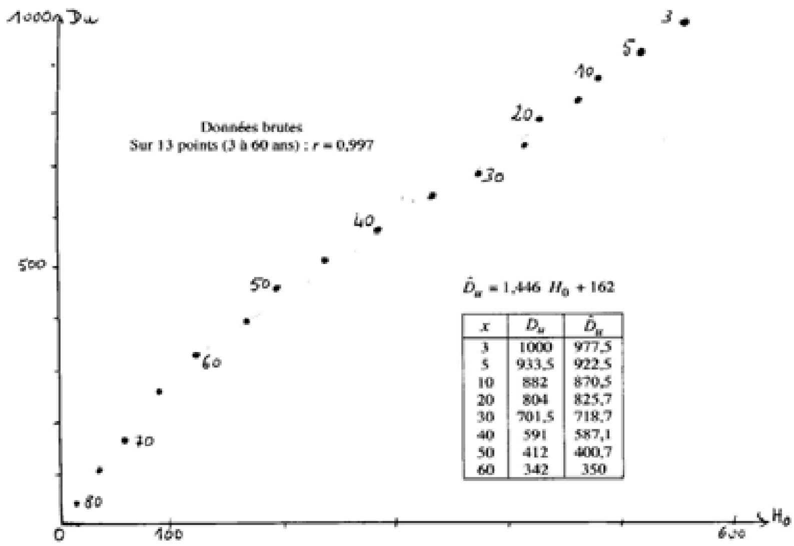
Sin embargo, hasta una veintena de años, dichas poblaciones prácticamente se confunden. Después se separan netamente y se ordenan de la forma que podíamos preveer: más supervivientes entre los « privilegiados » de la tabla Deparcieux, muchos menos para los londinenses de la tabla de Hogdson, mientras que la tabla « universal » de Duvillard ocupa una posición intermedia. Finalmente, todo el mundo se reune en la vejez extrema.

FIGURA 3. Curvas de mortalidad de *Hodgson, Duvillard* et *Deparcieux.* (Base de 1.000 vivos en tres años)
(rojo: Hogdson, azul: Duvillard, verde: Deparcieux)



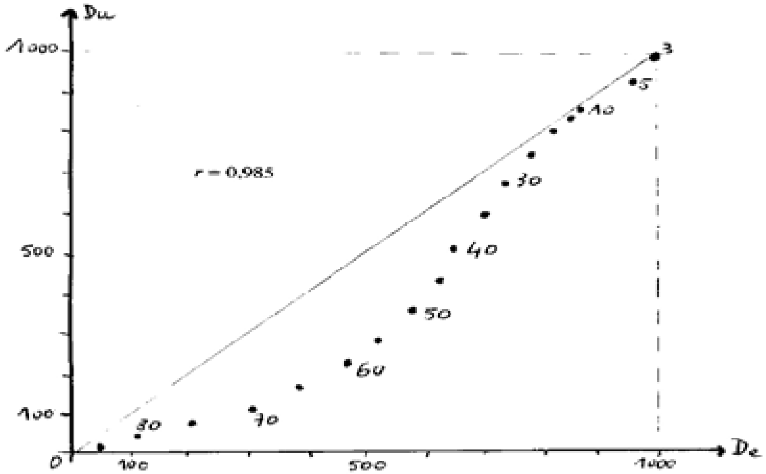
Pero veamos más de cerca. La Figura 4 revela una correlación muy fuerte entre la tabla de Hogdson y la de Duvillard, mientras que esta última se distingue mucho más netamente de la de Deparcieux a partir de 20 años de edad (Figuras 4 y 5).

FIGURA 4. Diagrama de correlación entre las tablas de *Hogdson* y de *Duvillard*. (Datos brutos – r : coeficiente de correlación lineal)



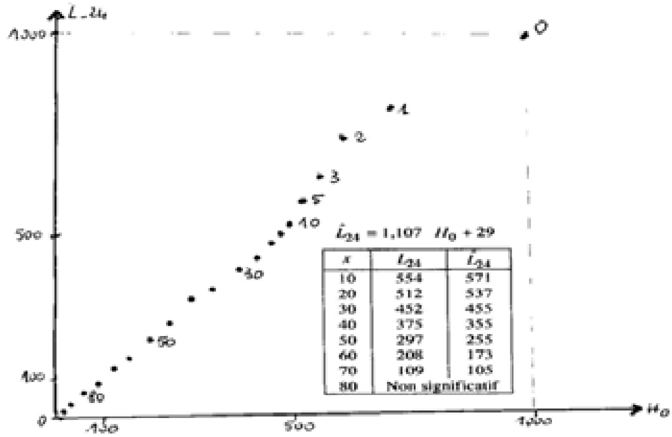
x	D_u	$\hat{D}_u$
3	1000	977,5
5	933,5	922,5
10	882	870,5
20	804	825,7
30	701,5	718,7
40	591	587,1
50	412	400,7
60	342	350

FIGURA 5. **Diagrama de correlación entre las tablas de *Deparcieux* y de *Duvillard*.**
(Base de 1.000 vivos en 3 años)



Por el contrario, la tabla Lambert 24 no se distingue netamente de la de Hogdson (Figura 6) salvo en las edades inferiores a cinco años, lo que significa una mortalidad infantil más leve en los datos que han servido para construir la primera, que en Londres en el decenio 1728-1737, que además difiere poco de la de los tres decenios 1728-1757 (cf. Cuadro VI y VII supra). Esto corrobora la hipótesis de que esta misteriosa tabla Lambert 24 concierne a una población universal (urbana y rural) y un período posterior.

FIGURA 6. **Diagrama de correlación entre las tablas de *Hogdson* y *Lambert 24*.**
(Base de 1.000 vivos al nacer)



Por último, las dos Figuras 7 y 8 ilustran gráficamente el carácter singular de la pirámide de edades en París en 1817 respecto a las tablas de mortalidad del siglo XVIII.

FIGURA 7. Curvas de mortalidad de *Fourier* y *Hogdson*
(Base de 1.000 vivos al nacer)
(en negro: J. Fourier, París 1817 ; en rojo Hogdson)
[Fourier : París intra muros, menos los militares, más los hospicios extra muros]

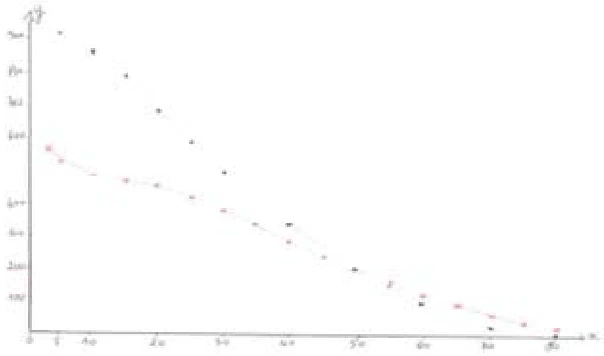
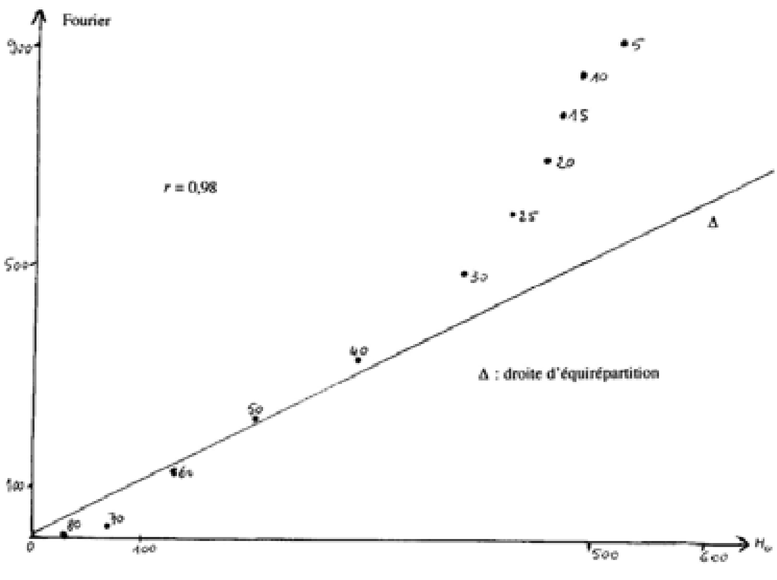


FIGURA8. Diagrama de correlación entre *Hogdson* y *Fourier*



N.B. Todos los coeficientes de correlación lineal son superiores a 0,97 incluso en los casos más desfavorables (Fourier respecto a Deparcieux, por ejemplo). Los valores tan elevados de ciertos coeficientes deben ser considerados en relación a este « estiaje ».

Además, cuando una de las variables es función monótona de la otra (lo que necesariamente es el caso aquí), es muy probable que su coeficiente de correlación sea muy elevado.

Sobre todo, como veremos en § 6, esas correlaciones lineales, casi perfectas, entre las tablas están relacionadas con el hecho de que cada una de nuestras tablas se ajusta muy bien a dos clases de leyes: el modelo lineal, y la ley de Gompertz. Esto explica aquello.

Destaca especialmente que el punto de inflexión, es decir, edad a la que la disminución del número de vivos es máxima, se sitúa entre 15 y 20 años para Fourier y a más de 40 años para Hogdson. Es 63 años para Lambert 24, 67 para Duvillard e incluso ¡73 años para Deparcieux !

4.3- A vueltas con los dos grupos « EL COMÚN DE LOS MORTALES » y los « PRIVILEGIADOS »

La Figura 9 muestra una correlación lineal muy fuerte entre la tabla de Halley y la de Duvillard. Hay también una correlación fuerte entre Halley y Lambert 24 o Hogdson, del mismo grupo, « el común de los mortales », tal y como se muestra en § 4.1.

Por el contrario, y sin que nos sorprenda, el diagrama de correlación entre Halley y Deparcieux (privilegiados) tendría el mismo aspecto que el de la Figura 5 más arriba.

Si ahora miramos el diagrama de correlación entre las tablas de Deparcieux y de Beauvisage (Figura 10) que conciernen a los «privilegiados», pero con más de un siglo de separación entre los períodos de observación, vemos un perfecto ajuste lineal.

FIGURA 9. **Diagrama de correlación entre las tablas *Halley* y *Duvillard***

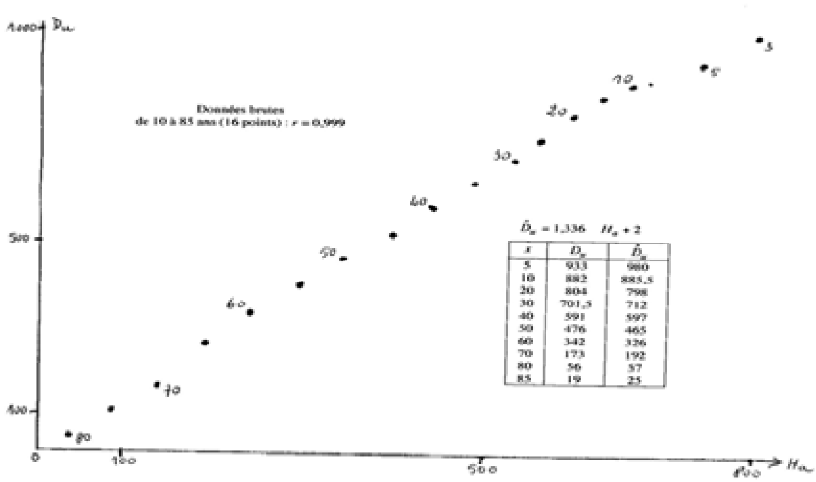
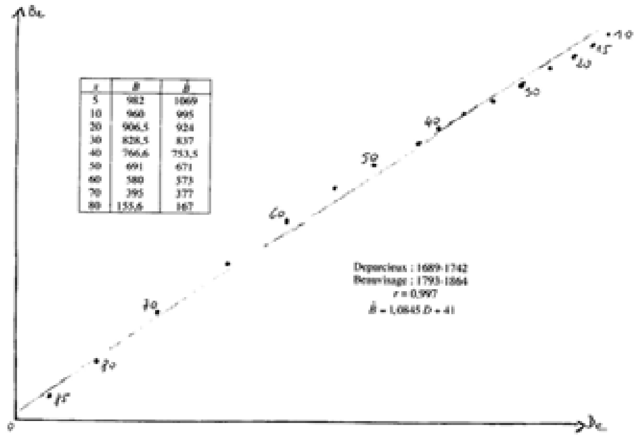
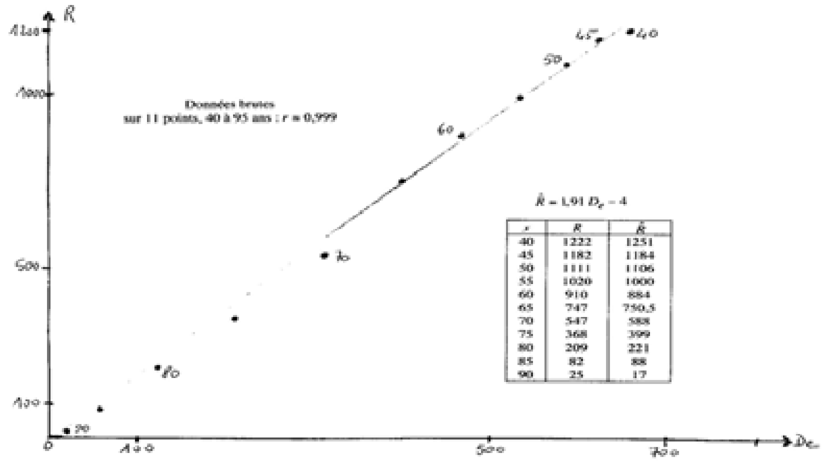


FIGURA 10. Diagrama de correlación entre tablas de *Deparcieux* et *Beauvisage*
(Base de 1.000 vivos en 3 años)



Más sorprendente, la casi perfecta correlación lineal entre la tabla de Deparcieux y la « micro-tabla » de las 1222 personalidades del Robert (Figura 11) de edades de 40 años o más de edad.

FIGURA 11. Diagrama de correlación entre *Deparcieux* y *Robert*



Pero aún más sorprendente resulta la muy buena correlación entre la tabla de Deparcieux y la de los 232 aristócratas de la micro-tabla Choisy, de 40 o más años de edad.

El diagrama de correlación es también casi tan rectilíneo como los dos anteriores.

Aquí me bastará presentar los resultados bajo la forma del Cuadro VIII más abajo. Los cálculos han sido efectuados sobre 12 puntos (con edades tomadas de 5 en 5, de 40 a 95 años) y tomando los datos brutos.

CUADRO VIII

Correlación entre *Deparcieux* y *Choisy*

$r = 0{,}995 \qquad \hat{C} = 0{,}3644\ \mathrm{De} + 8$

X	40	45	50	55	60	65	70	75	80	85	90
Choisy	232	227	217	206	189	168	125	82	53	21	8
$\hat{C}$	247,4	234,5	219,7	199,7	177	152	121	85	51	25,5	12

Estas últimas correlaciones, no sólo legitiman el uso de nuestras micro-tablas para las comparaciones, sino que muestran además que para construir su tabla, Deparcieux habría podido contentarse con una muestra bien distribuida en el tiempo y del orden del millar de observaciones.

Ciertamente, se comprende que los actuarios necesiten mucha más precisión. Pero para la comparación que me interesa aquí (esperanza de vida, número de supervivientes), lo que cuenta es el orden de magnitud. Lo que ponen de manifiesto los Cuadros VI a VII' es bastante significativo para que resulte útil comentarlo más a fondo.

5- Leyes teóricas de la mortalidad

Desde el siglo XVII, se han propuesto o utilizado algunas « fórmulas » que expresan el número de vivos y(x) en edad x en función explícita de x.

La mayor parte de estas leyes se expresan suponiendo la variable edad x continua e y(x) derivable. Se sitúan en la hipótesis de estabilidad.

En esos casos, la mortalidad m(x) se expresa de esta forma (siendo y(x) decreciente, su derivada es negativa):

(5.1) $\quad m(x) = -\,y'(x) = |y'(x)|,$

que sustituye a la expresión (1.1) del tiempo discreto. Igualmente, (1.2) y (1.3) se transforman en:

(5.2) $\quad y(x) = -\int_x^{\infty} y'(t)\,dt$

(5.3) $\quad Y(x) = \int_x^{\infty} y(t)\,dt,$

de donde en cuanto a la vida y la supervivencia media en edad x:

(5.4) $\quad v(x) = x + s(x) = x + \dfrac{1}{y(x)} \int_x^{\infty} y(t)\,dt = x + \dfrac{Y(x)}{y(x)}.$

Por lo que se refiere a la tasa de mortalidad a la edad x, será evidentemente:

(5.5) $\quad \tau(x) = -\dfrac{y'(x)}{y(x)} = \dfrac{|y'(x)|}{y(x)}.$

5.1- Tres Leyes de los siglos XVII y XVIII

Primera ley: J. Graunt, 1662 (cf. [1]) estima que el número de supervivientes disminuye en una relación de 5/8 cada diez años.

Expresado matemáticamente (Graunt no lo hizo) esto significa que y(x) es una función exponencial de forma:

(5.1.1) $y(x) = k\, e^{-\lambda x}$ $(\lambda > 0)$, donde e es la base de los logaritmos naturales con $e^{-10\lambda} = \frac{5}{8}$, es decir $\lambda \simeq \frac{1}{21,3}$.

Pero en este caso, tendríamos:

$$\tau(x) = \frac{|y'|}{y} = \lambda \quad \text{y} \quad s(x) = e^{\lambda x}\int_x^\infty e^{-\lambda t}\,dt = \frac{1}{\lambda}.$$

La tasa de mortalidad sería la misma en todas las edades, y la duración media de la vida sería constante: ¿sería entonces la vida humana eterna? Según el valor de λ dado por Graunt, a cada edad nos quedaría aún como promedio aproximadamente 21 años más de vida.

Es evidente que este « modelo » es totalmente irreal.

Segunda ley †: G.W. Leibniz hacia 1680 y A. de Moivre en 1725 (cf. [2] y [9]) supusieron que es la propia mortalidad la que tiene un valor constante a; de donde:

(5.1.2) $y(x) = a(c - x)$,

aquí c es la edad máxima teórica.

De hecho, esta linealidad de y(x) es una aproximación bastante buena de las curvas empíricas, no sobre la duración total de la vida, sino en un intervalo conveniente, acotado por las dos inflexiones de la curva (cf. Figuras 1, 3 y 7 supra); ya hemos hecho la observación a propósito de la tabla Lambert 24. Veremos ejemplos de ajuste lineal en § 6.

N.B. Como por otra parte, las curvas de mortalidad son muy cercanas en las primeras edades y en la vejez, esto explica las fuertes correlaciones lineales de § 4: dos funciones lineales de una misma variable son, en efecto, función lineal una de la otra.

Tercera ley: En 1772, J.H. Lambert propone (cf. [6] y [7]) la fórmula:

(5.1.3) $$y(x) = A\left(\frac{c-x}{c}\right)^2 - B\left(e^{-\lambda x} - e^{-\mu x}\right) .$$

Esta vez, hay una teoría explicitada. Es evidente que A es un parámetro de escala, el número de vivos y(0) al nacer tomado como base; en este caso, A = 10.000; c es la edad máxima aproximada.

† Agradezco a I. Schneider que me haya confirmado que de Moivre utilizaba la ley lineal. Lo que era dudoso al leer el « Discours préliminaire » de los padres Gaeta y Fontana en su traducción del anexo a la « Doctrine of chances » [traducida al francés por Pierre Crépel] en [8], p. 259 y ss.

El término del cuadrado $\left(\frac{c-x}{c}\right)^2$ significa, según Lambert, que la vida humana fluye así como el agua lo hace de un recipiente agujereado en su base: ese agua experimenta la ley de la gravedad, y su altura en el recipiente disminuye en la misma proporción que el cuadrado del tiempo (aquí, x).

Cada uno de nosotros puede además observar como se vacía su bañera: muy rápidamente cuando la bañera está llena, después mucho más lentamente.

En cuanto al término complementario $B\left(e^{-\lambda x}-e^{-\mu x}\right)$, Lambert lo justifica a través de algunas analogías con su teoría del calor, muy poco convincentes bajo mi punto de vista.

De hecho, para estimar los parámetros de su modelo se ha apoyado en la « Tabula X » de Süssmilch (cf. Anexo III, infra) ; ésta se detalla en [7]. En esta ocasión, él da como valores de los parámetros:

$$c=96, \qquad B=6176, \qquad \lambda=\frac{1}{13{,}682}, \qquad \mu=\frac{1}{2{,}23114}.$$

¡Admirable precisión ! Con estos valores para los parámetros, la ley Lambert teórica (cf. Anexo II) se tabula facilmente, y es – salvo para la edad en torno a los 3 años –notablemente cercana a la tabla J. P. (Londres, 1728-1757), que se resume en la « Tabula X », así como a la de Hogdson (Londres, 1728-1737).

Por desgracia esta ley no se aplica bien a ninguna de las otras tablas que hemos estudiado. Y lo mismo pasa en Lambert 24, ¡que el propio Lambert publicó en [6] !

5.2- **La Ley de B. Gompertz**, 1825 (cf. [11])

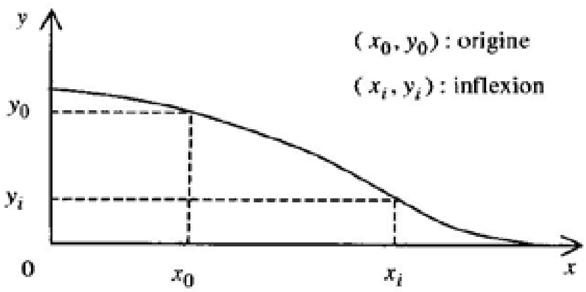
La teoría elaborada por Benjamin Gompertz es simple y se enuncia así : à partir de una cierta edad, la tasa de mortalidad $\tau(x)$ crece exponencialmente con la edad x.

Traducción matématica:

(5.2.0) $\quad \frac{y'}{y}=\alpha\, e^{\mu x}$, de donde

(5.2.1) $\quad$ y(x) = k $e^{-u(x-x_0)}$ con u(x) = $\frac{1}{\lambda}e^{\mu x}$, $\quad \lambda\,,\ \mu>0$.

FIGURA 12



La curva representativa tiene el aspecto que se muestra aquí arriba (cf. Figura 12). Para representar la totalidad de una curva de mortalidad del siglo XVIII, necesitaría una segunda inflexión, entre 0 y x_0 (cf. Figura 1, por ejemplo).

La inflexión x_i es la solución de: u(x - x_0) = 1.

Los parámetros se interpretan fácilmente. Tenemos primero:

(5.2.1)′ $$k = y(x_0)\, e^{1/\lambda} = y(x_i)\, e \quad \text{de donde} \quad \lambda = \left[1 - \ell_n\left(\frac{y_0}{y_i}\right)\right]^{-1},$$

x_0 es la edad elegida como origen; habitualmente se toma 40 años como valor de este origen, pero no es obligatorio.

A continuación se muestra por el cálculo que:

(5.2.2) $$\mu = \tau(x_i) = \frac{|y'(x_i)|}{y(x_i)},$$

μ es así la tasa de mortalidad para la edad en la que la mortalidad es máxima.

En cuanto a λ, tenemos:

(5.2.3) $$\lambda = \tau(x_i)/\tau(x_0) = \frac{|y'(x_i)|}{y(x_i)} : \frac{|y'(x_0)|}{y(x_0)},$$

λ es la razón entre la tasa de mortalidad en la inflexión y la tasa de mortalidad en el origen.

El valor de μ depende sólo de la unidad de tiempo elegida para la edad x, mientras que λ depende del valor elegido para el origen.

Además, se deduce de (5.2.1), (5.2.2) y (5.2.3) que λ y μ están vinculados por:

(5.2.4) $$\lambda = e^{\mu(x_i - x_0)},$$

lo que equivale a:

(5.2.4)′ $$x_i = x_0 + \frac{1}{\mu}\,\ell_n(\lambda),$$ donde ℓ_n es el logaritmo neperiano (o natural).

Así mismo, se deduce de (5.2.1) :

- la vida probable o la vida mediana p(x) en edad x

(5.2.5) $$p(x) = \frac{1}{\mu}\ell_n\left(\lambda\, \ell_n(2) + e^{\mu(x - x_0)}\right) + x_0,$$

- la supervivencia media s(x) a la edad x

(5.2.6) $\quad s(x) = \frac{1}{\mu}\Phi\left[u(x - x_0)\right],$ donde

(5.2.7) $\quad \Phi(z) = e^z \int_z^{\infty} \frac{e^{-t}}{t}\, dt,\quad z > 0.$

Podemos por tanto demostrar que p et s son, a edad constante, funciones crecientes de λ, y decrecientes de μ. Concretamente, tenemos en el origen:

(5.2.8) $\quad p(x_0) = \frac{1}{\mu}\ell_n(1 + \lambda\, \ell_n(2)) + x_0 \quad$ y $\quad s(x_0) = \frac{1}{\mu}\Phi\left(\frac{1}{\lambda}\right).$

N.B. La función Φ y sus derivadas satisfacen para todo $z > 0$ las inecuaciones:

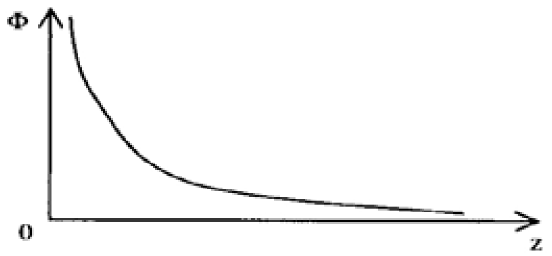
(5.2.9) $\quad \frac{1}{z} - \frac{1}{z^2} < \Phi\,(z) < \frac{1}{z},\quad \Phi'(z) = \Phi(z) - \frac{1}{z}$

y

$$\Phi''(z) = \Phi(z) - \frac{1}{z} + \frac{1}{z^2}.$$

Estas relaciones demuestran que la curva representativa de Φ tiene el aspecto de la Figura 13 (por otra parte, se puede además demostrar que: $\lim_{z\to 0} \Phi\,(z) = \infty$).

FIGURA 13



Por medio de su definición (5.2.7), podemos calcular los valores aproximados de Φ. Teniendo en cuenta el método que he utilizado (método de trapecios), esas aproximaciones son mejores cuanto mayor es z (cf. Cuadro IX).

CUADRO IX

Valores aproximados de Φ

z	$\Phi(z)$
0,1	2,04
0,2	1,5
0,3	1,225
0,4	1,05
0,5	0,925
0,6	0,83
0,7	0,75
0,8	0,69
0,9	0,638
1	0,597
1,5	0,458
2	0,362

Z	$\Phi(z)$
2,5	0,309
3	0,263
3,5	0,235
4	0,208
4,5	0,189
5	0,170
6	0,145
7	0,126
8	0,112
9	0,099
10	0,08745
11	0,0705

5.3- Métodos de ajuste a una Ley de Gompertz

El ajuste a una ley de Leibniz de una tabla empírica no plantea ningún problema, ya que se trata de un ajuste lineal clásico; sólo hay que elegir previamente un intervalo de edades adecuado sobre las que realizar dicho ajuste.

En lo que concierne a la ley de Gompertz, hay por lo menos cuatro métodos posibles.

- Método de inflexión. Teniendo en cuenta la interpretación (5.2.2) de μ, localizamos la edad x_i de mortalidad máxima, y la mortalidad $m(x_i)$ correspondiente. Así, se estima μ por :

(5.3.1) $$\hat{\mu} = \frac{m(x_i)}{y(x_i)} = \tau(x_i).$$

Por ejemplo, para Lambert 24 (cf. Anexo I), tenemos $x_i = 63$, $m(x_i) = 105$, $y(x_i) = 1772$. De donde:

$$\hat{\mu} = \frac{105}{1772} \simeq \frac{1}{16{,}9} \simeq 0{,}0592\,.$$

A continuación intentamos determinar un origen x_0 por tanteo (y por consiguiente λ), satisfaciendo la ecuación (5.2.3) más arriba, es decir:

(5.3.2) $$\frac{y(x_0)}{m(x_0)} e^{\hat{\mu} x_0} = \frac{y(x_i)}{m(x_i)} e^{\hat{\mu} x_i} = 707{,}7 \qquad \text{(para Lambert 24)}$$

Por medio de la tabla, obtenemos así:

x	… 40	… 48	49	50…
$\frac{y}{m} e^{\hat{\mu} x}$	498	694	709	714,5

Lo que plantea: $x_0 = 49$. De donde:

$$\lambda = e^{\hat{\mu}(x_i - x_0)} = e^{\frac{14}{16{,}9}} \simeq 2{,}29\,.$$

Obtenemos finalmente la estimación de k tomando como ejemplo la media de las que resultarían de (5.2.1)’, sea:

(5.3.3) $$\hat{k} = \frac{1}{2}\left(y_0\, e^{1/\lambda} + y_i\, e\right) = \frac{1}{2}(4711 + 4819) = 4765\,.$$

N.B. Por supuesto, la ecuación (5.3.2), tendría siempre la solución « trivial »: $x_0 = x_i$, y $\lambda = 1$.

También se podría estimar λ mediante $\left[1 - \ell_n\left(\frac{y_0}{y_i}\right)\right]^{-1}$. En este caso, daría: $\hat{\lambda} = 2{,}28$.

Este método, que tiene la ventaja de participar del espíritu de los de Lambert (véase, [6] y [7]), a menudo no da más que la solución $x_0 = x_i$. En general produce peores estimaciones que los otros.

- Método de las tasas de mortalidad. Se plantea w(x) = $-\ \ell_n(\tau(x)) = -\ \ell_n\left(\frac{|y'|}{y}\right)$.

Teniendo en cuenta la definición (5.2.1), resulta:

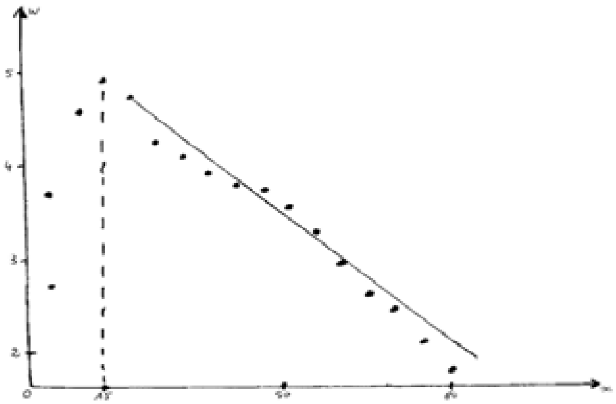
(5.3.4) $\qquad$ $w(x) = -\ \mu\, x + \ell_n\left(\frac{\lambda}{\mu}\right)$.

En esta fórmula, x es la edad relacionada con el origen elegido. En la aplicación: x = edad – 40.

Por lo tanto, a partir de una cierta edad, la cuestión se reduce a un problema de ajuste lineal.

A partir de la tabla, obtenemos las relaciones empíricas $\frac{y(x)}{m(x)}$, y calculamos las $\hat{w}(x) = \ell_n\left(\frac{y(x)}{m(x)}\right)$. Para la tabla Lambert 24 tenemos así, gráficamente, los puntos de la Figura 14.

FIGURA 14



Vemos que están aproximadamente alineados a partir de la edad de 15 años. Tomando como origen $x_0 = 40$, obtenemos, sobre los 11 puntos de 25 a 75 años, la línea recta de regresión $\hat{w}$ (con un coeficiente de correlación mediocre: 98 %).

$$\hat{w} = -\ \frac{x}{24{,}3} + 3{,}82,$$

de donde las estimaciones: $\hat{\mu} = \frac{1}{24{,}3} \simeq 0'041$, $\hat{\lambda} = 1{,}88$, $\hat{k} = 6360$ et $\hat{x}_i = 55{,}5$.

Estas estimaciones se encuentran muy lejos de las obtenidas por otros métodos. No obstante, el lector podrá verificar que el ajuste obtenido es muy bueno para las edades entre 10 y 30 años.

- Método del logaritmo iterado. Elegimos de una vez por todas un intervalo h; de nuevo, según la definición (5.2.1) la ley de Gompertz implica que:

(5.3.5) $$W(x) = \ell_n\left(\ell_n\left(\frac{y(x)}{y(x+h)}\right)\right) = \mu\, x + \ell_n\left(\frac{e^{\mu\, h}-1}{\lambda}\right).$$

En (5.3.5) x es la edad relacionada con el origen elegido; en general: x = edad - 40

Sea cual sea el resultado de (5.3.5), el problema se reduce de nuevo a un ajuste lineal.

En el caso de Lambert 24, con un intervalo de 5 años (h = 5) y sobre los 9 puntos, tomados de 5 en 5 años, de 40 a 80 años, obtenemos así la línea recta de regresión:

$$\hat{W} = \frac{x}{18} - 2{,}363,$$

con un muy buen coeficiente de correlación. De donde se obtienen las estimaciones:

$$\hat{\mu} = \frac{1}{18}, \quad \hat{\lambda} = \left(e^{\frac{5}{18}} - 1\right) e^{2{,}363} = 3{,}5, \quad \hat{k} = y(40)^{\ 1/\hat{\lambda}} = 4970, \quad \hat{x}_i = \frac{1}{\hat{\mu}}\ell_n(\hat{\lambda}) + 40 = 62{,}55.$$

La concordancia entre la edad estimada de inflexión, $\hat{x}_i$ con la edad observada (63 años), conduce a pensar que, en el caso de Lambert 24 este método va a dar un buen ajuste. El lector puede verificarlo; pero aquí el ajuste ya no es bueno para las edades más jóvenes, sino para el intervalo 30 – 80.

- Método mixto « inflexión-logaritmo iterado»

Si se plantea y/k = z, tenemos, según (5.2.1):

(5.3.6) $$Z(x) = \ell_n\left(\ell_n\left(\frac{1}{z}\right)\right) = \mu x - \ell_n(\lambda).$$

Sea, ateniéndose todavía a (5.2.1) :

(5.3.6)′ $$Z(x) = \ell_n(1 + \ell_n(y_i) - \ell_n(y)) = \mu x - \ell_n(\lambda).$$

En el caso de Lambert 24, tenemos: $y_i = 1773$, y en consecuencia:

(5.3.6)″ $$Z(x) = \ell_n(8{,}48043 - \ell_n(y(x)) = \mu x - \ell_n(\lambda).$$

De donde, por ajuste lineal, y tomando la base $x_0 = 40$:

$$\hat{\mu} = 1/17{,}3373, \quad \hat{\lambda} = 3{,}7654 \quad \text{y} \quad \hat{k} = e\, y_i = 4819{,}5.$$

N.B. El lector podrá verificar que este ajuste es, para el intervalo de 20 a 85 años, tan bueno, incluso mejor, que el que se obtiene por el tercer método. No obstante, este tercer método es más conocido y se emplea más.
Ello se debe a que el punto de inflexión de la curva de mortalidad, por lo menos en las tablas del siglo XVIII, no suele estar tan bien marcado como en la tabla Lambert 24 ; hay una cierta indeterminación respecto al valor de y_i ; ahora bien, es indispensable conocer este valor para aplicar este método, como también lo es conocer el método de « la inflexión » (1er método).

5.4- A título de información, La Ley de Makeham, 1860

Al igual que Gompertz, Makeham asume que la tasa teórica de mortalidad es función exponencial de la edad, pero le añade un « índice mínimo » ν:

(5.4.0) $$-\frac{y'}{y} = \frac{|y'|}{y} = \nu + \alpha\, e^{\mu x}.$$

De donde la expresión de y(x) :

(5.4.1) $$y(x) = k \exp\left[-(\nu x + \frac{1}{\lambda}\, e^{\mu x}\right].$$

N.B. De ahora en adelante, la notación exp(z) significa e^z.

Al constar de un parámetro más que el de Gompertz, este modelo se ajusta generalmente mejor a los datos empíricos.

Pero a mí me parece que no hay una interpretación clara de los parámetros (en concreto, ν sería la tasa de mortalidad a una edad infinita negativa, lo cual no tiene ningún sentido). Por tanto, una condición necesaria de validez de un modelo es que todos sus parámetros sean interpretables en la realidad que se supone que representa.

6.- Ajustes a las Leyes de Gompertz y Leibniz

Vamos ahora a dar los resultados de ajuste de algunas tablas a la ley de Gompertz por un lado, y a la de Leibniz por el otro.

Subrayemos que el ajuste a una ley de Gompertz es muy bueno para todas y no sólo para aquellas que son objeto de los cuadros que siguen.

Los Cuadros XV y XVI al final del párrafo recapitulan los valores estimados de los parámetros para el conjunto de las once tablas estudiadas en este artículo.

CUADRO X

Ajustes de la Tabla de Halley

Leibniz : $\hat{y}(x) = 9{,}09\ (86{,}5 - x)$ Gompertz : $\hat{y}(x) = 781\ \exp\left(-\frac{1}{3{,}742}\, e^{\frac{x-20}{27}}\right)$.

	Leibniz		Halley	Gompertz				
x	%	$\hat{y}$	y	$\hat{y}$	%	σ	$\hat{\sigma}$	x
3	4,9	759	798	677				3
5	1,2	740,5	732	670				5
7	4,5	722,5	692	622	4,4			7
10	5,2	695	661	649,4	1,7	40,4	42,1	10
15	3,3	649,5	628	625,5	0,4	37,5	38,2	15
20	1	604,2	598	597,9	0,02	34,2	34,5	20
25	1,5	559	567	566,2	0,15	30,9	31	25
30	3,4	513	531	530,35	0,1	27,9	27,7	30
35	4,5	468	490	490,2	0,04	24,1	24,6	35
40	5,1	422,5	445	445,9	0,2	22,3	21,7	40
45	5	377,5	397	397,8	0,2	19,1	19	45
50	4,2	331,5	346	346,8	0,2	17,25	16,7	50
55	2	286,3	292	294	0,7	14,5	14,5	55
60	0,6	240,6	242	241	0,4	12,4	12,5	60
69	4,6	152	159					
70	5,4	142,7	142	142,3	0,2	7,6	9,2	70
72			120	124,9	4			72
80		58,8	41	66,3				
85								

N.B. Las líneas de puntos ------ indican los límites de las edades cuya aproximación es ≤ 5 %.

Notaciones para los Cuadros X a XIV
x la edad
y numero de supervivientes
$\hat{y}$ estimación de y
s supervivencia media observada
$\hat{s}$ supervivencia media estimada
σ supervivencia probable observada
$\hat{\sigma}$ supervivencia probable estimada
% **diferencia relativa** $\frac{|y-\hat{y}|}{y}$, en porcentaje
Para el cálculo de $\hat{s}$ y $\hat{\sigma}$, véanse las fórmulas (5.2.5), (5.2.6) y el Cuadro IX (§ 5.2) más arriba

Los ajustes de la Tabla de Hogdson son algo peores que los de la Tabla de Halley.

CUADRO XI

Ajustes de la Tabla de Hogdson

Leibniz : $\hat{y}(x) = 8\ (77{,}5 - x)$

Gompertz : $\hat{y}(x) = 800\ \exp\ . \left(- e^{\frac{x-40}{33}} \right).$

	Leibniz		Hogdson	Gompertz				
x	%	$\hat{y}$	y	$\hat{y}$	%	s	$\hat{s}$	x
5	10,3	580	526	566	7	39,3	37,8	5
10	10,2	540	490	535	9	37	34,65	10
15	5,3	500	475	501	5,7	33,15	31,75	15
20	0,2	460	459	464	1,1	29,2	29	20
25	1,4	420	426	424	0,5	26,25	26,45	25
30	1,3	380	385	382	0,8	23,75	24	30
35	0	340	340	339	0,3	21,5	21,75	35
40	2	300	294	294,5	0,2	19,7	19,55	40
45	5,7	260	246	250	1,6	17,9	18,2	45
50	8	220	204	206,5	1,25	16,1	16,5	50
55	9	180	165	165,5	0,3	14,3	14,65	55
60	7,7	140	130	128	1,55	12,5	13	60
65	1	100	99	95	4	10,6	11,5	65
70	13	60	69	67	2,9	9,15	1045	70
75		20	45	44,5	1,1	7,75	9	75
80			29	27,8	4,15	5,7	8	80
85			14	16	14,3			85
90			5	8,5				90
95			0	4				95

N.B. Inflexión teórica: $\hat{x}_i = 40$ y $\hat{\tau}_i = \mu = 1/33$.

Inflexión observada: $x_i = 40{,}5$ y $\tau_i = 10/290 = 1/29$.

Constatamos, sin que nos sorprenda, que la ley teórica de Lambert (cf. Anexo II) es muy cercana a la de Hogdson y se ajusta también muy bien a las dos leyes, la de Leibniz y la de Gompertz.

CUADRO XII

Ajustes de la tabla *Lambert theórica*

Leibniz: $\hat{y}(x) = 8,62\ (76 - x)$.

Gompertz: $\hat{y}(x) = 834,5\ 834,5\ \exp\left(- e^{\frac{x-40}{31,8}}\right)$.

	Leibniz		Lambert théorica	Gompertz		
x	%	$\hat{y}$	y	$\hat{y}$	%	x
5		612	549			5
10	10,5	569	515	559		10
15	3,2	523	507	523	3,2	15
20	0,3	482,7	484	484,9	0,2	20
25	0,1	439,6	440	443,6	0,9	25
30	1,6	396,5	403,5	400	0,9	30
35	0,75	353,4	356	354	0,6	35
40	1	310,3	307	307	0	40
45	3,15	267,2	259	260	0,4	45
50	4,9	224,1	213,6	214	0,2	50
55	5,7	181	171,3	170,3	0,6	55
60	3,8	138	133	130,4	2	60
65	4,25	94,8	99	95,4	4,6	65
70		51,7	69,6	66,2	4,8	70
75		8,62	45,3	43,2	4,9	75
80			26	26,2	0,8	80
85			12	14,6	22	85
90			3,4	7,35		90

N.B. Inflexión teórica: $\hat{x}_i = 40,7$ y $\hat{\tau}_i = \mu = 1/31,8$.
Inflexión observada: $x_i = 40$ y $\tau_i = 98/3110 = 1/31,75$.

En cuanto a la Tabla Fourier del censo de la población Parísina de 1817, una vez más, se distingue netamente de las otras por dos características: los ajustes son aceptables sólo para los primeros decenios de la vida (de 0 a 40 años); y el ajuste a la ley de Gompertz es claramente peor que en las otras tablas.

CUADRO XIII

Ajustes de la Tabla *Fourier* (París, 1817)

Leibniz : $\hat{y}(x) = 16\ (63{,}25 - x)$.

Gompertz : $\hat{y}(x) = 1400\ \exp\left(\frac{-1}{0{,}7}\ e^{\frac{x-40}{28}}\right)$.

	Leibniz		Fourier	Gompertz		
x	%	$\hat{y}$	y	$\hat{y}$	%	x
0	1	1010	1000	998	0,2	0
5	0,15	930	932	933,6	0,2	5
10	2	850	867	861,9	0,55	10
15	3,2	770	795,5	783,2	1,55	15
20	0,35	690	692,5	698,5	0,85	20
25	2,2	610	597	609	2	25
30	5,5	530	502	516,75	2,9	30
40	8,2	370	342	335	2	40
50	0,7	210	211,5	180,3	14,6	50
60		50	105,5	74,4	29,3	60
70			33	20,85	57	70
80			5	3,4	31	80
90			0,2	0,25	20	90

N.B. Inflexión teórica: $\hat{x}_i = 30$.

Inflexión observada: $15 < x_i < 20$ ($x_i = 20 - \varepsilon$), $\varepsilon > 0$.

Finalmente se puede leer aquí abajo el ajuste de la micro-tabla Robert a una ley de Gompertz (cf. Anexo IV). El ajuste a una función lineal tendría poco sentido, ya que se refiere a una población muy reducida en edades por debajo de los treinta años.

Recordemos que la micro-tabla Choisy y su ajuste a la ley de Gompertz figuran en el Anexo V.

CUADRO XIV

Ajuste de la Tabla *Robert* a une ley de Gompertz

$$\hat{y}(x) = 1330 \ \exp\left(-\frac{1}{12}\, e^{x/13}\right).$$

X	y	$\hat{y}$	$\hat{y}-y$	%	σ	$\hat{\sigma}$	x
15	1295	1314	19	1,45			
20	1293	1306	13	1	47,5	47,65	20
25	1288	1295	7	0,55	42,5	43,15	25
30	1277	1279	2	0,15	38	38,55	30
35	1256,5	1256,5	0	0	33,5	33,25	35
40	1225	1224	- 1	0,08	28,5	29,15	40
45	1176,5	1177	0,5	0,04	24,5	24,8	45
50	1109	1112	3	0,25	20,5	20,65	50
55	1018,5	1023	4,5	0,45	16	16,85	55
60	902	905	3	0,33	12,5	14,4	60
65	747,5	756	8,5	1,15	9,7	10,3	65
70	559	581	22	4	7,5	7,9	70
75	368,5	394,5	26	7	6	5,9	75
80	212	224	12	5,7	4	4,3	80
85	97	97,6	0,6	0,6	3	3,1	85
90	30,8	28,8	2	6,7	2	2,2	90

N.B. Aquí los datos brutos (Anexo IV) han sido ajustados por medio de una media móvil de 11 puntos.

Recapitulaciones

Modelo de Leibniz: el cuadro XV recapitula los datos relativos al ajuste a esta ley para ocho tablas (recordemos que las Tablas J. P. y Lambert teórica son casi iguales).

Me parece importante remarcar tres observaciones:

- la alta calidad de los ajustes (coeficiente r) que se confirma en los cuadros X a XIII

N.B. Ésta muestra muy claramente que, como ya se ha señalado al final de § 4.2, las fuertes correlaciones lineales que hemos constatado entre las tablas reflejan la calidad de los ajustes de estas tablas a la ley Leibniz.

En efecto, si tenemos aproximadamente:

(6.1) $y_1 = a_1 x + b_1$ y $y_2 = a_2 x + b_2$, con $a_1 a_2 \neq 0$, entonces tenemos más o menos:

(6.2) $y_2 = \frac{1}{a_1}\,(a_2 y_1 + a_1 b_2 - a_2 b_1)$.

- el crecimiento del parámetro a (mortalidad anual media para el tramo de edades considerado) es conforme con lo que era de esperar;
- la originalidad aberrante de la Tabla Fourier (París, 1817). Lo cual no es ninguna novedad.

CUADRO XV

Ajuste a la ley de Leibniz : $y = a(c - x)$.

Tablas	Fechas	*I*	*n*	*r*	*a*	*I’*
Beauvisage	1793-1864	15-65	11	0,988	6,3	
Deparcieux	1689-1742	15-65	11	0,994	6,75	4 - 63
Duvillard	Avant 1789	15-65	11	0,998	7,2	10 - 62
Halley	1678-1691	35-80	10	0,999	7,8	25 - 80
Hogdson	1728-1737	20-70	11	0,998	8,1	16 - 54
Lambert 24	Avant 1770	25-80	12	0,999	8,44	21 - 67
Lambert théorique	1728-1757	20-70	11	0,998	8,6	15 - 65
Fourier	1817	0-60	10	0,997	16	0 - 30

N.B. Todos los datos están elaborados en relación a una base de 1000 al nacer**.**

Notaciones.

I : intervalo sobre el que se realiza el ajuste
n : número de puntos utilizados (generalmente, en intervalos de 5 años)
r : coeficiente de correlación lineal
a : coeficiente de regresión = mortalidad media anual
I’ : intervalo de diferencia relativa ≤ 5 % entre los datos y el modelo

Modelo de Gompertz: el parámetro más importante de este modelo es μ, la tasa de mortalidad en edad x_i (inflexión) de mortalidad máxima. Por otra parte, recordemos (cf. § 5.2) que la supervivencia media y la supervivencia probable son funciones decrecientes de μ.

Respecto al Cuadro XVI, hay tres observaciones dignas de destacar también:

- encontramos de nuevo los mismos grupos y subgrupos que en § 4.1 (cf. Cuadros VI a VII’): privilegiados, «el común de los mortales» universal, « el común de los mortales » urbano;
- la casi concordancia entre los valores estimados y los valores observados para los dos parámetros escogidos. Por ello me parece inútil sobrecargar este cuadro con los datos de los coeficientes de correlación, que son todos aproximadamente del 0,99;
- aquí todavía Fourier se distingue de las otras Tablas. Las tasas μ et τ_i son de la misma magnitud que en el último grupo, mientras que $\hat{x}_i$ y x_i se separan notablemente, y difieren mucho una de la otra.

CUADRO XVI

Ajustes a la ley de Gompertz : $y(x) = k \exp\left(-\frac{1}{\lambda} e^{\frac{x-x_0}{\mu}}\right)$.

Tablas	μ	$\hat{x}_i$	τ_i	x_i	I'
Choisy	1/11,16	76,5			20 – 85 *
Beauvisage	1/12,75	73,3	1/11,95	74,5	25 - 65
Robert	1/13	72,3			15 - 75
Deparcieux	1/14,5	71	1/13	73	30 - 70
Duvillard	1/17,4	63,5	1/14,5	67,5	31 - 80
Lambert 24	1/17,9	62,5	1/16,9	63	35 - 80
Halley	1/27	55,5	1/25	59	10 - 70
Lambert théorique	1/31,8	40,7	1/3175	39,5	13 - 82
Hogdson	1/33	40	1/28	40,5	15 - 60
Fourier	1/28	30	1/25	≃ 20	0-30

N.B. Salvo en el caso de *Halley* $(x_0 = 20)$, se ha tomado en todo momento 40 años como la edad de origen.

Notaciones.

$\mu = \hat{\tau}_i$ = tasas de mortalidad teórica en la inflexión

$\hat{x}_i$ = inflexión estimada = $\frac{1}{\mu} \ell_n (\lambda) + x_0$

τ_i = tasas de mortalidad observada en la inflexión observada

x_i = inflexión observada

I' = intervalo de diferencia relativa ≤ 5 %

* excepto cerca de 70 años

N.B. El hecho de que todas las tablas se ajusten muy bien a una ley de Gompertz explica a su vez las correlaciones observadas entre las tablas.

En efecto, la función $y(x) = k \exp\left(-\frac{1}{\lambda} e^{\mu x}\right)$ genera una serie entera.

Cerca del punto (x_0, y_0), los primeros términos de esa serie son:

$$(6.3) \quad \frac{y}{y_0} = 1 - \frac{\mu}{\lambda}(x - x_0) - \frac{\mu^2(\lambda - 1)}{2\lambda^2}(x - x_0)^2 - \frac{\mu^3(\lambda^2 - 3\lambda + 1)}{6\lambda^2}(x - x_0)^3 (1 + \varepsilon(x - x_0)), \text{ con: } \lim_{z \to 0} \varepsilon(z) = 0.$$

El cuadro XVI aquí debajo y los valores obtenidos para λ en los ejemplos tratados, muestran que los coeficientes de los términos de grado 2 y 3 son netamente inferiores a 10^{-2} y 10^{-3} respectivamente. Resulta que cerca del punto (x_0, y_0), la curva de Gompertz se aproxima muy bien a la tangente de ese punto, es decir, la línea recta de la ecuación:

$$(6.4) \quad \frac{y}{y_0} = 1 - \frac{\mu}{\lambda}(x - x_0).$$

En particular, si tomamos como origen el punto de inflexión (x_i, y_i), tenemos, $\lambda = 1$; el término cuadrático desaparece, como es debido, y (6.3) pasa a:

$$(6.3)' \quad \frac{y}{y_i} = 1 - \mu(x - x_i) + \frac{\mu^3}{6}(x - x_i)^3(1 + \varepsilon(x - x_i)).$$

Cerca de la inflexión, la curva es aproximada al 3^er grado, es decir, muy aproximada a la tangente:

(6.4)′ $\frac{y}{y_i} = 1 - \mu(x - x_i)$.

Así nos encontramos prácticamente como en el caso de los ajustes de las tablas de mortalidad a una función lineal (ley de Leibniz).

Se podría llegar a la misma conclusión partiendo no de la función misma y(x), sino de una fórumula de linealización como la de más arriba (5.3.6).

Si $y_1(x)$ y $y_2(x)$ son dos funciones de mortalidad Gompertz, y si se plantea:

$$Z(x) = \ell_n\left(\ell_n\left(\frac{k}{y(x)}\right)\right).$$

Se deduce de (5.3.6) la relación lineal:

(6.5) $\mu_2 Z_1 - \mu_1 Z_2 = \mu_1 \ell_n(\lambda_2) - \mu_2 \ell_n(\lambda_1)$.

Si se supone, a fin de aligerar la escritura, $\lambda_1 = \lambda_2 = 1$ (es decir, que el origen en ambas funciones es su inflexión), obtenemos:

(6.5)′ $Z_1 = \frac{\mu_1}{\mu_2} Z_2 = \nu Z_2$ $(\nu > 0)$.

de donde:

$$e^{Z_1} = \ell_n\left(\frac{k_1}{y_1}\right) = \left(e^{Z_2}\right)^{\nu} = \left(\ell_n\left(\frac{k_2}{y_2}\right)\right)^{\nu},$$

Y finalmente:

(6.6) $y_1 = k_1 \exp\left[-\left(\ell_n\left(\frac{k_2}{y_2}\right)\right)^{\nu}\right]$.

Si se plantea $\eta = 1 - \frac{y_2}{k_2}$ para facilitar el cálculo, el desarrollo en serie de Taylor del segundo miembro de (6.6) resulta como aproximación de primer grado de y_1 en función de y_2 la relación lineal:

(6.7) $y_1 \simeq \frac{k_1\ \mu_2}{k_2\ \mu_1} y_2$.

7.- Algunas observaciones a modo de conclusión

Esto es el trabajo de un aficionado que no está especializado en demografía histórica, ni está familiarizado a fortiori con sus métodos. Es probable que para los especialistas no aporte nada nuevo, nada que no sea ya bien conocido y destacado.

No obstante, voy a correr el riesgo de sacar tres conclusiones, por muy triviales que puedan parecer.

1. La profunda igualdad entre las tablas de mortalidad de la época estudiada, igualdad de la que son testigo por una parte las fuertes correlaciones entre ellas y por otra la calidad de sus ajustes a una misma familia de leyes teóricas – las leyes de Gompertz y, en menor grado, las leyes de Leibniz.

2. Si hay igualdad, también hay diferencias. En nuestra época son numerosos los estudios de lo que podríamos llamar la « mortalidad diferencial »: diferencias entre clases o categorías sociales, entre zonas geográficas, entre modos de vida, etc. ¿Y el pasado, y especialmente, el siglo dieciocho? Aquí he puesto de manifiesto, de manera esbozada, diferencias entre « privilegiados » y el resto, entre poblaciones urbanas y poblaciones no diferenciadas (seguro que esto también es « descubrir el Mediterráneo »). Seguramente es posible ir mucho más allá.

3. Ahora bien, para hacer esto, no es necesario disponer de datos relativos a poblaciones numerosas, como sucede con los mormones.

De hecho, pequeñas muestras del orden del millar de individuos o menos, han sido suficientes para construir tablas muy adecuadas. Ahora bien, hay múltiples fuentes análogas a las que yo he utilizado y son accesibles a todos y cada uno de nosotros.

Por ejemplo, es posible individualizar algunas categorías dentro de muestras escogidas en compilaciones como los diccionarios (pero hay muchas otras fuentes): los artistas especialmente, pero también los escritores, o incluso los científicos.

Me ha impresionado especialmente que al hacer con el Journal de los Goncourt lo mismo que había hecho (y presentado aquí) con las Mémoires del Abate de Choisy, resulta que, aunque se trate del siglo diecinueve, la longevidad es sensiblemente menor que en mis tablas de Robert y Choisy o en la de Deparcieux.

Esto probablemente se deba a que las referencias de los hermanos E. y J. de Goncourt eran relativas a periodistas, escritores y artistas.

Me parece que la principal dificultad en este tipo de trabajo, realizado sin medios informáticos, radica en evitar las repeticiones. Afortunadamente, los investigadores actuales están más avanzados que yo y saben utilizar los odenadores.

En suma, aún queda mucho trabajo que hacer para quienes se sientan tentados a « lanzarse » a esa tarea.

ANNEXE I

Table de mortalité de Lambert, tirée d'une table de Süßmilch

Zum §. 24. der Anmerkungen über die Sterblichkeit.

Alter Jahre.	Jährlich Sterbende.	Lebende.	Summe der Lebenden.	Es stirbt einer von	Mittleres Alter.	Alter, wo die Hälfte gestorben.
0	2610	10000	295022	4	29,5	22,6
1	610	7390	285022	12	40,2	40,5
2	340	6780	277632	20	42,1	44,4
3	233	6440	270852	27	45,0	46,7
4	169	6207	264412	37	46,7	48,3
5	140	6038	258205	43	47,7	49,5
6	116	5898	252167	51	48,5	50,2
7	96	5782	246269	60	49,6	51,0
8	80	5686	240487	71	50,3	51,6
9	68	5606	234801	82	50,9	52,0
10	58	5538	229195	95	51,4	52,4
11	50	5480	223657	110	51,8	52,8
12	45	5430	218177	121	52,2	53,1
13	42	5385	212747	128	52,5	53,3
14	39	5343	207362	137	52,8	53,6
15	37	5304	202019	143	53,1	53,8
16	35	5267	196715	151	53,3	54,0
17	33	5232	191448	159	53,6	54,2
18	36	5199	186216	144	53,8	54,4
19	39	5163	181017	132	54,1	54,6
20	43	5124	175854	119	54,3	54,8
21	48	5081	170730	106	54,6	55,1
22	54	5033	165649	93	54,5	55,3
23	58	4979	160616	86	55,2	55,6
24	62	4921	155637	79	55,6	55,9
25	65	4859	150716	75	56,0	56,3
26	67	4794	145857	72	56,4	56,7
27	69	4727	141063	69	56,8	57,0
28	70	4658	136336	67	57,3	57,4
29	71	4588	131678	65	57,7	57,8
30	72	4517	127090	63	58,1	58,1
31	74	4445	122573	60	58,6	58,5
32	75	4371	118128	58	59,0	58,9
33	77	4296	113757	56	59,5	59,3
34	79	4219	109461	54	59,9	59,7
35	80	4140	105242	52	60,4	60,1
36	80	4060	101102	51	60,9	60,5
37	81	3980	97042	49	61,4	60,9
38	81	3899	93062	48	61,9	61,3
39	80	3818	89163	48	62,4	61,7
40	80	3738	85345	47	62,8	62,1
41	79	3658	81607	46	63,3	62,5
42	78	3579	77949	45	63,8	62,9
43	77	3501	74370	45	64,2	63,2
44	76	3424	70869	45	64,7	63,6
45	75	3348	67445	44	65,1	64,0
46	75	3273	64097	44	65,5	64,3
47	76	3198	60924	42	66,0	64,7
48	77	3122	57626	41	66,4	65,0
49	78	3045	54504	39	66,9	65,4
50	80	2967	51459	37	67,3	65,8
51	82	2887	48492	35	67,8	66,2
52	84	2805	45605	33	68,2	66,6
53	86	2721	42800	32	68,7	67,0
54	88	2635	40079	30	69,2	67,5
55	90	2547	37444	28	69,7	67,9
55	90	2547	37444	28	69,7	67,9
56	92	2457	34897	27	70,2	68,4
57	94	2365	32440	25	70,7	68,9
58	96	2271	30075	24	71,2	69,4
59	98	2175	27804	22	71,8	70,0
60	100	2077	25629	21	72,3	70,6
61	101	1977	23552	20	72,9	71,2
62	103	1876	21575	18	73,5	71,9
63	105	1773	19699	17	74,1	72,4
64	104	1668	17926	16	74,7	73,1
65	102	1564	16258	15	75,4	73,7
66	99	1462	14694	15	76,0	74,4
67	95	1363	13232	14	76,7	75,1
68	92	1268	11869	14	77,4	75,7
69	88	1176	10601	13	78,0	76,3
70	85	1088	9425	13	78,7	76,9
71	82	1003	8337	12	79,4	77,5
72	80	921	7334	12	80,0	78,1
73	79	841	6413	11	80,6	78,7
74	77	762	5572	10	81,3	79,3
75	75	685	4810	9	82,0	79,9
76	73	610	4125	8	82,7	80,6
77	70	537	3515	8	83,5	81,3
78	66	467	2978	7	84,3	82,0
79	62	401	2511	6	85,2	82,8
80	56	339	2110	6	86,2	83,6
81	50	283	1771	6	87,2	84,5
82	43	233	1488	5	88,4	85,7
83	35	190	1255	5	89,6	87,4
84	25	155	1065	6	90,8	89,3
85	18	130	910	7	92,0	90,9
86	13	112	780	9	92,9	92,0
87	10	99	668	10	93,7	92,8
88	9	89	569	10	94,4	93,0
89	9	80	480	9	95,0	94,1
90	8	72	400	9	95,5	94,7
91	8	64	328	8	95,1	95,3
92	8	56	264	7	96,7	95,9
93	7	48	208	7	97,3	96,5
94	7	41	160	6	97,9	97,1
95	7	34	119	5	98,5	97,7
96	6	27	85	4	99,2	98,3
97	6	21	58	3	99,8	98,9
98	5	15	37	3	100,5	99,5
99	4	10	22	2	101,2	100,3
100	3	6	12	2	102,0	101,0
101	1	3	6	2	102,0	102,5
102	1	2	3	2	103,5	103,0
103		1	1		104,0	

ANEXO II

La ley teórica de Lambert

Población y(x) a cada edad x; vida probable M(x) y supervivencia probable M(x) – x.

x	*y*	M(*x*)	M(*x*) - *x*
0	10 000	16,875	"
1	8 146	29,386	28,4
2	6 965	36	34
3	6 224	39,57	36,6
4	5 786	41,84	37,84
5	5 491	43,38	38,4
6	5 330	44,23	38,2
7	5 240	44,7	37,7
8	5 192	44,96	37
9	5 167	45,09	36,1
10	5 153,5	45,16	35,15
11	5 143,4	45,22	34,2
12	5 132,2	45,28	33,3
13	5 117	45,36	32,35
14	5 096,4	45,47	31,45
15	5 070	45,61	30,6
16	5 035,7	45,8	29,8
17	4 995,4	46	29
18	4 948,6	46,26	28,25
19	4 896	46,55	27,55
20	4 840	46,85	26,85
21	4 774,2	47,21	26,2
22	4 705,9	47,58	25,6
23	4 633,3	47,98	25
24	4 556,9	48,4	24,4
25	4 477	48,85	23,85
26	4 393,8	49,32	23,3
27	4 308	49,8	22,8
28	4 219,8	50,45	22,45
29	4 129,5	50,82	21,8
30	4 035,5	51,37	21,35
31	3 943,9	51,92	20,9
32	3 849	52,45	20,45
33	3 753,2	53	20
34	3 656,5	53,6	19,6
35	3 559,5	54,18	19,2
36	3 461,8	54,78	18,8
37	3 364	55,4	18,4
38	3 266	56	18
39	3 168,4	56,62	17,6
40	3 070	57,25	17,25
41	2 973,9	57,88	16,9
42	2 877,3	58,52	16,5
43	2 781,5	59,17	16,15
44	2 686,3	59,82	15,8
45	2 592	60,5	15,5
46	2 498,7	61,12	15,1
47	2 406,3	61,8	14,8

x	*y*	M(*x*)	M(*x*) - *x*
48	2 315	62,46	14,45
49	2 225	63,12	14,1
50	2 136	63,93	13,9
51	2 048,8	64,55	13,55
52	1 962,7	65,13	13,15
53	1 878	65,8	12,8
54	1 794,7	66,45	12,45
55	1 713	67,17	12,15
56	1 633	67,85	11,85
57	1 544,6	68,54	11,55
58	1 477,8	69,22	11,2
59	1 403	69,9	10,9
60	1 330	70,6	10,6
61	1 257,7	71,3	10,3
62	1 189	72	10
63	1 120	72,67	9,65
64	1 053,5	73,37	9,35
65	989,5	74,06	9
66	927	74,79	8,8
67	866,5	75,5	8,5
68	808	76,15	8,15
69	751	76,85	7,85
70	696	77,56	7,55
71	644	78,25	7,25
72	593	78,97	7
73	544	79,65	6,65
74	497,5	80,35	6,35
75	453	81,03	6
76	410	81,73	5,75
77	369,5	82,62	5,6
78	331	83,425	5,4
79	294,5	84,175	5,15
80	260	84,54	4,55
81	227,5	85,25	4,25
82	197	85,95	3,95
83	182	86,5	3,5
84	143	87,3	3,3
85	119	88	3
86	97	88,7	2,7
87	77,2	89,4	2,4
88	59,51	89,94	1,95
89	43,93	90,75	1,75
90	30,4	91,43	1,45
91	19,146	91,98	,9801
92	9,94	92,7	0,7
93	2,9	93,25	0,25
94	0	0	0

ANNEXE III

Table de Süßmilch des décès à Londres en 30 ans de 1728 à 1757
(*Die göttliche Ordnung*… II, Annexe, p. 31)

31

TABVLA X.

Liste der in London Gestorbenen nach den Jahren und zwar in 30 Jahren, vom Jahr 1728 bis 1757.

v. a Collection of the yearly Bills of mortality from 1657 to 1758 inclus. London. 1759. 4. p. 141.

Alter der Gestorbenen.	von 1728 bis 1732.	von 1733 bis 1737.	von 1738 bis 1742.	von 1743 bis 1747.	von 1748 bis 1752.	von 1753 bis 1757.	Summa aller 30 Jahr.
Unter 2 Jahr	50363	52796	49538	41948	39887	38371	272903
2 bis 5 Jahr	10984	12521	12540	9939	9069	9069	64745
5—10	4834	4941	4979	4292	3652	3223	25912
10—20	4267	3975	4546	3787	3374	2942	22891
20—30	10216	9560	11162	9920	9724	7892	58474
30—40	12271	12631	13729	11678	11816	9977	71502
40—50	12077	11912	14297	12659	12130	10163	73238
50—60	9754	9939	11914	9704	9543	8928	59782
60—70	8459	7850	8516	7661	7692	7091	47269
70—80	5553	5131	6793	6134	5043	5025	33679
80—90	3464	2986	3150	2598	2343	2407	16948
90 und drüber	671	595	565	433	352	363	2979
	132913	134237	142720	120753	114625	106074	750322

Unter jedem Tausend Gestorbener sind also gewesen

Alter.	erste 5 Jahr.	zweyte 5 Jahr.	dritte 5 Jahr.	vierte 5 Jahr.	fünfte 5 Jahr.	sechste 5 Jahr.	General-Verhältn.
Unter 2 Jahr	378	393	349	347	347	361	363
2—5	83	93	89	82	80	92	87
5—10	37	37	35	36	31	30	34
10—20	32	29	32	31	30	28	31
20—30	76	72	78	82	85	74	77
30—40	93	89	97	97	103	94	96
40—50	91	89	101	105	106	96	97
50—60	73	74	84	80	83	84	80
60—70	64	59	60	64	67	67	63
70—80	41	38	48	50	44	47	45
80—90	24	22	23	22	20	23	23
90 und drüber	8	5	4	4	4	4	4
	1000	1000	1000	1000	1000	1000	1000

ANEXO IV

x	m	y	Y	v	p
15	1	1295	65.441	65	67,75
16	-				
17	-				
18	-				
19	-				
20	1	1294	58.970	65,1	67,75
21	1	1293			
22	-				
23	-				
24	1	1292			
25	2	1291	53.799	66,2	67,8
26	2	1289			
27	-				
28	4	1287			
29	5	1283			
30	4	1278	47,362	66,5	68
31	3	1274			
32	4	1271			
33	2	1267			
34	3	1265			
35	4	1262	41.007	67	68,1
36	7	1258			
37	10	1251			
38	9	1241			
39	10	1232			
40	3	1222	34.763	68	68,5
41	9	1219			
42	4	1210			
43	16	1206			
44	8	1190			
45	12	1182	28.716	68,8	69
46	10	1170			
47	16	1160			
48	16	1144			
49	17	1128			
50	11	1111	22.932	70,15	69,8
51	18	1100			
52	24	1082			
53	15	1058			
54	23	1043			
55	19	1020	17.538	71,7	71

m	m	y	Y	v	p
55	19	1020	17.538	71,7	71
56	20	1001			
57	21	981			
58	24	960			
59	26	936			
60	35	910	12.640	73,4	72,6
61	27	875			
62	22	848			
63	29	826			
64	31	797			
65	32	766	8.384	75,45	74,4
66	45	734			
67	50	689			
68	48	639			
69	44	591			
70	38	547	4.965	78,6	77,7
71	35	509			
72	30	474			
73	44	444			
74	41	400			
75	35	359	2591	81,7	81
76	31	324			
77	27	293			
78	36	266			
79	21	230			
80	32	209	1119	84,85	84,1
81	26	177			
82	20	151			
83	19	131			
84	30	112			
85	17	82	339	88,65	88
86	15	65			
87	11	50			
88	10	39			
89	4	29			
90	9	25	74	92,45	92
91	3	16			
92	4	13			
93	3	9			
94	2	6			
95	3	4	5		

Diccionario « Robert des nombres propes». Personalidades europeas fallecidas por muerte natural entre 1667 y 1801.

x: edad $m(x)$: número de muertos a la edad x $y(x)$: número de vivos de edad x

$Y(x) = \sum_{i=0}^{\infty} y(x+i)$: número de vivos de edades $\geq x$ $v(x)$: vida probable a la edad x

p(x) : vida probable a la edad x

N.B. Una de estas personalidades vivió 100 años: Fontenelle, muerto en 1757

ANEXO V

Abate de Choisy, *Mémoires*
Aristócratas fallecidos entre 1623 y 1749
(Mismas notaciones que para el Anexo IV)

$\hat{y}$ y $\hat{p}$: valores calculados según el ajuste a una ley de Gompertz

$$\hat{y} = 242 \ \exp\left(-\frac{1}{23{,}3}\ e^{\frac{x-40}{11{,}16}}\right)$$

x	m	y	Y	V	p	$\hat{y}$	$\hat{p}$	$\frac{\lvert y-\hat{y}\rvert}{y}\%$
20		245	11.774	67,5	70,5	240,3	71,15	2
	4							
25		241	10.560	68,3	71	239,3	71,2	0,8
	3							
30		238	9.361	68,8	71	237,8	71,3	0
	4							
35		234	8.184	69,5	71,15	235	71,5	0,45
	2							
40		232	7.020	69,75	71,2	232	71,7	0
	5							
45		227	5.869	70,35	71,4	226	72	0,44
	10							
50		217	4.757	71,40	71,8	218,6	72,6	0,74
	11							
55		206	3.696	72,45	72,15	205,3	73,4	0,34
	17							
60		189	2.696	73,75	73,375	187	74,7	1
	21							
65		168	1.807	75,25	75,17	131,7	76,15	3,75
	43							
70		125	1.121	79,5	78,2	128,7	78,25	3
	43							
75		82	664	82,5	81,5	90	80,95	10
	29							
80		53	318	85,5	83,8	51,5	84,15	3
	32							
85		21	123	90,5	89	21,5	87,7	2,4
	13							
90		8	38			5,5		
	5							
95		3	6			0,65		
97	2							
		0						
100	0							
1	2	3	4	5	6	7	8	9

Tablas y curvas de mortalidad del S. XVIII
Semejenzas y divergencias

Bibliografia general

[1] GRAUNT J., *Observations naturelles et politiques*, Eric Vilquin (éd.), édition Institut National d'Études Démographiques (I.N.E.D.), París, 1977.

[2] *LEIBNIZ G.W. et les raisonnements sur la vie humaine*, J.-M. Rohrbasser et J. Véron (éds), édition Institut National d'Études Démographiques (I.N.E.D.), París, 2001.

[3] DEPARCIEUX A., *Essai sur les probabilités de la durée de la vie humaine* (1746), addition à l'Essai (1760), C. Behar (ed.), édition Institut National d'Études Démographiques (I.N.E.D.), París, 2003.

[4] *A collection of yearly bills of mortality from 1657 to 1758 inclusive,* printed for A. Millar in the Strand, London, 1759.

[5] SÜSSMILCH J.P., *L'ordre Divin*

1741, J.-M. Rohrbasser (éd.), édition Institut National d'Études Démographiques (I.N.E.D.), París 1998.

1761, J. Hecht (éd.), édition Institut National d'Études Démographiques (I.N.E.D.), París, 1979.

[6] LAMBERT J. H., *Contributions mathématiques à l'étude de la nuptialité et de la mortalité, 1765 et 1772*, J.-M. Rohrbasser et J. Véron (éds), édition Institut National d'Études Démographiques (I.N.E.D.), París, 2006.

[7] BARBUT M., ROHRBASSER J.-M., VERON J., "Lambert et la loi de survie", Mathématiques et Sciences humaines 171, 2005.

[8] *Arithmétique politique dans la France du XVIIIe siècle*, mélanges sous la direction de Th. Martin, édition Institut National d'Études Démographiques (I.N.E.D.), París, 2003.

[9] DE MOIVRE A., *Annuities upon lives*, London, 1725. Adjunto a The doctrine of chances (1735), edición de 1756.

[10] *Recherches statistiques sur París et le département de la Seine*, C. Ballard imprimeur du Roi, París, 1821.

[11] GOMPERTZ B., "On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies", Phil. Trans. Of the Royal Soc of London, vol. 115, 1825.

[12] THUILLIER G., *Le premier actuaire de France* : DUVILLARD (1755-1832), París, La Documentation française, 1997.

[13] BEAUVISAGE E., *Des tables de mortalité et de leurs applications aux assurances sur la vie,* París, Gauthier-Villars, 1867.

[14] THUILLIER G., *Une ténébreuse affaire : la caisse Lafarge (1787-1892)*, París, La documentation française, 1999.

[15] QUICQUET A., *Aperçu historique sur les formules d'interpolation des tables de survie et de mortalité*, París, Warnier & Cie, 1893.

Rubrica « Démographie mathématique »

Revista : *Mathématiques et Sciences humaines*
Mathematics and Social Sciences

http://msh.revues.org/

http://www.ehess.fr/revue-msh/index.php

n° 148, 1999 ; n° 149, 2000 ; n° 153, 2001 ; n° 156, 2001 ; n° 159, 2002 ; n° 160, 2002 ; n° 164, 2003 ; n° 166, 2004 ; n° 167, 2004 ; n° 171, 2005

Capítulo 10

Des plans d'expérience un siècle avant Ronald A. Fisher

ÉRIC BRIAN
Directeur d'Etudes a l'EHESS
(Centre Maurice-Halbwachs),
Associe a l'INED

Introduction

L'histoire des statistiques accorde à Ronald Fischer (1890-1962) la primeur en 1925 de la conception des plans d'expérience (*experimental design*), c'est-à-dire d'un dispositif expérimental de mise à l'épreuve de la variabilité statistique d'un phénomène. On sait que la période de l'entre-deux-guerres fut celle de la mise au point des tests inférentiels (Stigler, 1986 et 1999 ; Hald, 1998). Pourtant, on va le voir, on s'est préoccupé de variabilité et de ce qui ressemble bien à un plan d'expérience, cela pour des question d'agronomie et d'hérédité déjà un siècle plus tôt, bien avant que la mesure de la variance n'ait été mise au point. C'était un temps où le calcul numérique relevait aux yeux des naturalistes et de médecins de la science expérimentale (Schweber, 1995 ; 2006). Ajoutons que la postérité de l'artisan de ce calcul a été réservée à ses seules qualités d'homme de lettres… L'épisode – et son oubli – sont exemplaires[1].

Probabilité et transmission des qualités humaines à Paris à la fin du XVIIIe siècle

En 1788, la disparition du naturaliste Buffon a donné au Secrétaire perpétuel de l'Académie des sciences Condorcet, qui n'aimait guère ni la méthode ni le style de son aîné, l'occasion de

1. Cet article est un élément des travaux qui ont conduit à la publication de Brian et Jaisson (2007a et 2007b). Cette recherche a été soutenue par l'Institut national des études démographiques (Paris). Sur la question chevaline, elle a bénéficié des conseils de Daniel Roche.

mentionner les travaux de l'abbé Lazzaro Spallanzani récemment connus à Paris pour avoir su procéder à l'insémination artificielle d'une chienne au moyen d'une seringue (Spallanzani, 1785). Pour Condorcet, les expériences de l'abbé, comme celles de médecin suisse Albrecht von Haller, contredisaient les vues générales du naturaliste (Condorcet, 1788).

De sorte que, quelques années plus tard, dans l'Esquisse d'un tableau historique des progrès de l'esprit humain, Condorcet achèvera son célèbre prospectus sur l'évocation des résultats asymptotiques à attendre quant aux progrès de l'espèce humaine, et en particulier de ceux des recherches empiriques de physiologie comparée.

> *« [...] Les facultés physiques, la force, l'adresse, la finesse des sens ne sont-elles pas au nombre de ces qualités dont le perfectionnement individuel peut se transmettre? L'observation des diverses races d'animaux domestiques doit nous porter à le croire : et nous pourrons le confirmer par des observations directes faites sur l'espèce humaine. » (Condorcet, Esquisse..., 1795 [2004, p. 458])*

Le passage annonçait des développements du même auteur sur l'avenir de l'humanité rendus public au XXe siècle seulement. Les fragments de la dernière époque du *Tableau historique*, réservée à la prospective, et depuis 2004 tout à fait accessibles, permettent de mieux saisir ce qu'avait à l'esprit le mathématicien philosophe. Le *Fragment 9* (« Sur l'Atlantide ») dit comment les savants organisés de manière collective pourraient espérer combattre les préjugés de leur temps :

> « *Enfin, il est des tentatives auxquelles, soit par la nature même de l'objet, soit par sa petitesse apparente, soit par l'extrême incertitude du succès, un seul homme craint de se livrer, parce qu'il s'exposerait soit au ridicule, soit à une sorte de honte : et qui ne sait combien on les craint encore, même quand on sent que ces flétrissures ne peuvent être imprimées que par la main d'un préjugé méprisable, combien on redoute l'opinion de ceux même dont on dédaigne le plus la raison ? [...] Je placerai dans cette classe les recherches commencées par Spallanzani sur une génération en quelque sorte artificielle, ou celles des causes qui déterminent le sexe, soit dans les foetus des animaux vivipares, soit dans les germes des oeufs. »* (Condorcet, [Tableau historique. Fragment 9], 1795, [2004, p. 910]).

L'organisation et la coordination de la société savante sont indispensables, aux yeux de Condorcet, dès qu'il s'agit, comme dans le cas de la génération, d'affronter les préjugés et les superstitions du siècle. Un autre passage, le *Fragment 10* (dit « Sur les effets moraux et politiques de quelques découvertes... »), est plus précis encore et livre une longue discussion de cette question :

> *« Dans l'hypothèse que l'on eût un moyen de déterminer à volonté le sexe des enfants, du moins jusqu'à un certain point, en naîtrait-il une différence sensible dans le nombre des individus de chaque sexe au lieu de l'égalité presque entière qui existe aujourd'hui, quel serait alors le plus nombreux, et quels pourraient être sur l'ordre social les effets de cette disproportion ?* » (Condorcet, [Tableau historique. Fragment 10], 1922, [2004, p. 932]).

L'enjeu moral – et dans les mots de Condorcet, social – des recherches sur l'insémination artificielle et sur la physiologie de la génération est précisé un peu plus loin :

> *« Quoique l'expérience de Spallanzani, unique jusqu'ici, ne soit appuyée que par des analogies prises d'espèces trop éloignées de la nôtre, cependant on ne doit pas la reléguer*

absolument dans la classe des chimères. On peut donc se demander ce qui résulterait pour l'espèce humaine, pour les progrès de la Civilisation, de la connaissance d'un moyen de séparer la reproduction des individus, de la réunion intime des individus qui doivent y concourir et des plaisirs physiques ou moraux attachés à cette union, comme on sait séparer ces mêmes plaisirs de la reproduction des individus. » [id., ibid., p. 935].

L'arrière-plan théologique de la question

L'attention que Condorcet a porté à la physiologie de la génération et à ses incidences morales a une raison toute mathématique. Sa conception de la proportion des naissances des deux sexes répondait à deux autres connues à la fin du XVIIIe siècle. La première était très repandue et relevait de la physico-théologie à la manière du Pasteur Johan Peter Süssmilch. Selon des démarches tout à fait conforme aux normes empiriques du XVIIIe siècle (il s'agit bien d'une physique), mais tendue vers un objectif théologique, la physico-théologie tenait la régularité du surcroît des naissances masculines (disons 105 garçons pour 100 filles) comme une preuve empirique de l'action d'une providence divine orienté vers la préservation de la monogamie, étant entendu que le constat de la plus forte mortalité des garçons avant l'âge du mariage ne faisait aucun doute : il fallait bien qu'il y ait plus de garçons à la naissance, pour qu'il y en ait autant au moment du mariage (Süssmilch, 1741). Il faut ici constater que la quasi-totalité des savants qui sont soucié de la question pendant la seconde moitié du XVIIIe siècle se sont partagé entre l'option physico-théologique et une option tout à fait sceptique à l'égard de l'emploi du calcul dans ces matières (le cas le plus connu fut D'Alembert).

La seconde conception savante du rapport des sexes à la naissance n'était maîtrisée que par quelques mathématiciens connaisseurs du calcul analytique des probabilités à la manière de Laplace, c'est-à-dire très peu de gens. Elle renvoyait le phénomène à un produit de pur hasard gouverné par une probabilité abstraite susceptible d'être estimée (Brian et Jaisson, 2007a ; 2007b). C'était comme tirage des boules blanches ou noires placées dans une urne, mais pas en nombre égal. Il y avait une cause inaccessible : « la plus grande facilité des garçons » (ou des boules de telle couleur) ; et cette cause avait un effet mesurable : la plus grande fréquence des garçons parmi les naissances. Ici il n'était pas nécessaire de faire l'hypothèse ni de la monogamie ni de la Providence. On se souvient de ce que Laplace répondra quelques années plus tard à Napoléon à ce sujet : Dieu ? Il n'avait pas besoin de cette hypothèse. Que le calcul des probabilités au début du XIXe siècle ait senti le soufre : il ne faut pas en douter. Laplace lui-même, dans l'*Essai philosophique* (1814) ne précisait-il pas « En détruisant la superstition, je suis loin de vouloir ébranler la religion » (Laplace, [1921, II, p. 63]).

Cette anecdote napoléonienne nous convainc que l'agnosticisme laplacien avait de quoi choquer ses contemporains. Son principal concurrent en mathématique à la veille de la Révolution, celui qui non seulement avait su saisir la puissance de ses calcul mais aussi les réorganisé dans un édifice de « métaphysique du calcul » (nous dirions aujourd'hui d'épistémologie des mathématiques) laissé à l'état d'ébauche, c'était précisément Condorcet. Et celui-ci considérait que Laplace accordait trop à cette mystérieuses « plus grande facilité des garçons ». Cette métaphysique que Laplace mettait à la porte de ses démonstrations mathématiques, Condorcet craignait qu'elle ne s'insinuât par la fenêtre du nouveau calcul des probabilités. Rien, si ce n'est la superstition et ses préjugés n'interdisait d'imaginer que l'humanité « eût un moyen de déterminer à volonté le sexe des enfants, du moins jusqu'à un certain point » (Condorcet), c'est-à-dire eu moyen d'intervenir sur cette « plus grande facilité des garçons » (Laplace). Les chiffres démontrent une régularité, le calcul laplacien fait comprendre que la causalité physico-théologique n'est pas justifiée, mais rien dans ce calcul n'indique ni pourquoi ni comment les naissances des garçons seraient « plus facile » (Brian et Jaisson, 2007a ; 2007b).

Il est temps d'observer que les manuscrits du *Fragment 10* sur la génération ne furent connus qu'au XXe siècle. Je viens d'indiquer le cheminement intellectuel de Condorcet, mais le XIXe siècle n'a disposé que d'une ébauche de quelques lignes à la fin de la Xe époque de l'*Esquisse* :

> « […] *Les facultés physiques, la force, l'adresse, la finesse des sens ne sont-elles pas au nombre de ces qualités dont le perfectionnement individuel peut se transmettre ? L'observation des diverses races d'animaux domestiques doit nous porter à le croire : et nous pourrons le confirmer par des observations directes faites sur l'espèce humaine*. » (Condorcet, *Esquisse...*, 1795 [2004, p. 458].

La physiologie est indiquée, la transmission des qualités physiques… Le périmètre des recherches qui nous intéressent est tracé. Il sera aussi celui de R. A. Fisher.

Des réactions à Londres dès 1798

On le sait, l'*Esquisse* de Condorcet fut l'une des principales cibles de Thomas Robert Malthus, en 1798, lors de la première publication de son *Essai sur le principe de population*. Lecteur du prospectus de la X^{e} Epoque du *Tableau historique*, le pasteur anglican avait rejeté l'idée que l'humanité puisse attendre une amélioration de son sort du fait de l'accroissement de la population et de celle de la durée de la vie humaine. En effet pour Malthus, l'accroissement de la population était en progression géométrique et celui des ressources en progression arithmétique. La double logique conduisait chez lui à l'anticipation d'une catastrophe que seule aurait pu faire éviter une limitation de l'accroissement de la population, différentiée selon les classes sociales. Malthus, en 1798, c'est-à-dire aux heures les plus noires de l'esprit public anglais, avait voulu stigmatiser la confiance dans les progrès de l'esprit humain plaidée par Condorcet, proclamer à nouveau la Révélation et, par voie de conséquence, affirmer que « les stimulations dues aux besoins intellectuels sont perpétuellement entretenues par la variété infinie de la nature et les ténèbres qui enveloppent les questions métaphysiques » (Malthus, 1798 [1980, p. 73-85 et 151-166]). L'envol mystique auquel aboutissait l'*Essai* de 1798 répondait ainsi à la méditation sur la perfectibilité asymptotique de l'espèce humaine et de son esprit, l'évocation qui avait ponctué l'*Esquisse* de Condorcet quelques années plus tôt. Une telle réaction obscurantiste, les Fragments du *Tableau historique* comme son *Esquisse* suggèrent que Condorcet, s'il l'avait observée, y aurait reconnu la marque de « la main d'un préjugé méprisable ».

La noirceur du pasteur a été critiquée par des auteurs anglais post-malthusiens. Parmi eux, une génération plus tard, on compte Michael Thomas Sadler, un parlementaire *Tory*, c'est-à-dire attaché à une conception politique d'inspiration religieuse, conservatrice et anti-libérale – et donc peu porté à revenir à Condorcet. Activiste du mouvement des *Poor Law* à la manière d'un christianisme social, il a publié, en 1830, *The Law of Population* (Sadler, 1830). L'ouvrage conjuguait, à l'encontre de la conjecture de la « surfécondité », deux conceptions de la population jusque là distinctes : l'une dans le sillage des deux éditions de l'essai de Malthus (1798 et 1803), mais contre son pessimisme, et l'autre issue de la lecture de Süssmilch. Or, pour un providentialiste, le différentiel des sexes à la naissance était induit par une anticipation divine des pertes masculines de telle sorte que la monogamie y retrouvât son compte. De là, chez Sadler, un chapitre intitulé *Of the law of population: anticipatory computations of Nature, especially in reference to the proportion of the sexes*.

Sadler, qui disposait des enregistrements familiaux d'une population de près d'un millier de Pairs du Royaume, avait scruté les proportions des naissances des deux sexes selon les âges absolus des parents, leurs âges relatifs, et – dispositif expérimental peu ordinaire conçu pour circonscrire la capacité d'anticipation de la Nature – les cas de remariages après veuvage. Le statisticien se garde bien d'indiquer à son lecteur l'intention politique et théologique du réformateur *Tory* qui, porté par la Révolution industrielle et animé par les débats parlementaires autour de la *Poor Law*, entendait quant à lui préserver l'action divine qu'il n'envisageait que sous un jour favorable. « Nous avons d'abondantes preuves du dessein bienveillant de la Déité » écrivait-il en concluant son chapitre sur les calculs anticipés.

Retour à Paris, en 1814 : guerres napoléoniennes et cheptels.

Entre temps, en France, les agronomes et les physiologistes avaient approfondi des recherches sur la génération des animaux domestiques, notamment celle des moutons, des vaches et des chevaux. Ce regain d'intérêt soutenu par le gouvernement sous le I^er^ Empire doit être rapporté au contexte des guerres européennes du début du XIX^e^ siècle, consommatrices à l'excès de chevaux pour la Cavalerie, le Train et les Etats-Majors, de laine et de cuir pour les vêtements et les couchages, de viande enfin pour le ravitaillement. Si bien qu'il faudrait mesurer les effets de ces bouleversements continentaux sur les variétés régionales des espèces domestiques. Il est ainsi raisonnable de considérer que la plupart des « races locales » encouragées au XIX^e^ siècle furent de pures inventions induites par ce cataclysme continental.

Parmi les savants très actifs dans ce domaine, on compte, aux abords de la X^e^ Section de la Première Classe de l'Institut, « Economie rurale et art vétérinaire »[2], deux expérimentateurs chevronnés, l'un monarchiste établi non loin de Paris, Charles-Gilbert Terray de Morel-Vindé[3], l'autre bonapartiste et installé en Aveyron, Charles Girou de Buzareingues[4]. L'un et l'autre se sont rencontrés dans le cercle de l'académicien vétérinaire Jean-Baptiste Huzard. Leurs travaux sur les moutons mérinos y trouvèrent quelque émulation. Les deux auteurs visaient une même cible, le savoir des éleveurs des campagnes, dont ils critiquaient les adages par des observations systématiques et commentées. Ils ont trouvé à l'Académie des sciences – où tous deux furent correspondants et l'un d'entre eux, membre – une chambre d'enregistrement de leurs constats et un lieu de consécration de leur science agronomique.

2. La Première Classe de l'Institut provient en grande part de l'ancienne Académie royale des sciences et reprendra cette dénomination à partir de 1816.

3. Né en 1759 et mort en 1842, Morel-Vindé était magistrat à la fin de l'Ancien régime, il s'est retiré de la vie publique pour se consacrer à l'agronomie pendant la Révolution. Continûment monarchiste constitutionnel, il fut correspondant de la Section d'Economie rurale de la Première Classe de l'Institut dès 1808, puis l'un de ses membres dans l'Académie des sciences de 1824, et d'autre part l'un des membres de la Société d'agriculture. Pair de France en 1815, il est entré au Conseil supérieur de l'agriculture en 1818. Il pratiquait l'élevage dans sa propriété de La Celle Saint-Cloud (voir Girardin, 1845).

4. Né en 1773 et mort en 1856, Girou a soutenu l'Empire jusque dans les Cent Jours, il fut correspondant de la Section d'Economie rurale en 1826 et membre de la Société d'agriculture. Il a échoué en décembre 1828 dans sa tentative pour succéder au fauteuil de Bosc à l'Académie des sciences, connaissant les entraves de la section d'Economie rurale. Voir à ce sujet Institut de France, Académie des sciences, *Procès-verbaux des séances*, Hendaye, Imprimerie d'Abbadia, 1921, t. IX, p. 149 et 151 (dans la suite *PV*) et Duval, 1858.

Morel-Vindé a publié en plusieurs livraisons, de 1813 à 1816, ses « Observations sur la monte et l'agnelage » dans les *Annales d'agriculture française*. Chaque année, il donnait un registre détaillé des naissances dans son élevage de moutons et des caractéristiques de ses reproducteurs (Morel-Vindé, 1813-1816). Dans le but de ruiner les lieux communs sur l'influence de l'âge des brebis sur le sexe de leur descendance, il a ainsi livré des tableaux de la répartition des naissances pour les années 1812 et 1813 en distinguant, année par année, l'âge des mères et les deux sexes des produits. Le propre de son étude, menée dans le même cadre pendant plusieurs années, fut d'organiser un numérotage de chaque animal et une collection de bulletins historiques associés à chaque numéro. Morel-Vindé disait prendre modèle sur les haras. Le bénéfice de la méthode était de pouvoir récapituler des états annuels, individus par individus, et d'en tirer des tableaux qui permettaient de conclure sur telle ou telle conjecture. Dans l'article de 1814, et pour la première fois, une nomenclature d'âge (d'année en année à partir d'un an et demi construite à des fins expérimentales à partir d'un registre *ad hoc* des mères) était croisée avec une répartition par sexe des naissances.

Notons que Morel-Vindé a publié à Paris, un avant Sadler à Londres, un écrit anti-malthusien (Morel-Vindé, 1829), plus précisément une critique des thèses des économistes politiques français qui, pendant les années 1820, revendiquaient la doctrine de Malhus. Il s'y affichait en défenseur du *laisser faire* – la loi française assurant la liberté de la propriété territoriale – convaincu de ce que les conclusions du pasteur ne pouvaient convenir qu'à l'Angleterre et à l'Irlande du tournant du siècle, et non pas à la France de la Restauration dont la population avait été suffisamment saignée par les guerres napoléoniennes pour qu'une politique d'inspiration malthusienne lui soit épargnée. La brochure ne dit rien sur les thèmes abordés par Sadler à Londres au même moment, ni sur les expériences ovines de La Celle Saint-Cloud. Sadler et Morel-Vindé, on le voit, écrivaient dans des univers différents.

Une théorie de la génération à Paris vers 1825

C'est dans une tout autre revue, les *Annales des sciences naturelles*, que Charles Girou de Buzareingues a fait paraître des « Observations sur les rapports de la mère et du père avec les produits, relativement au sexe et à la ressemblance » (Girou de Buzareingues, 1825). Il reprenait au passage les compilations de Morel-Vindé pour en tirer trois tableaux annuels de répartition du sexe des produits des agnelages selon l'âge des brebis. Girou de Buzareingues a proposé à cette occasion de distinguer trois cas, celui des brebis d'âge « moyen » (on dirait aujourd'hui plus précisément d'âge *médian*), et ceux des mères plus âgées ou plus jeunes. Nous récapitulons ces chiffres dans une même table cumulée.

	Agnelages 1812, 1813 et 1814			
Âge des brebis mères	mâles	femelles	totaux	% mâles
5 ans et ½ et plus	144	110	254	56,7%
4 ans et ½	68	69	137	49,6%
3 ans et ½ et moins	203	178	381	53,3%
Total	415	357	772	53,8%

Sans prendre particulièrement garde à la faiblesse des effectifs de la catégorie médiane, ni au fait que les deux sexes, en général, ne sont pas nécessairement à parité, Girou de Buzareingues en concluait, à l'encontre de son prédécesseur, que ses registres et ses chiffres mettaient en évidence un effet de l'âge de la mère sur le sexe de la descendance.

Or, les *Annales des sciences naturelles* étaient une revue nouvelle, créée un an plus tôt, et animée par une rédaction dynamique tenue par trois beaux-frères héritiers d'une dynastie savante active depuis la fin du XVIII[e] siècle et surtout sous le I[er] Empire : Jean Audouin, Adolphe Brongniart et Jean-Baptiste Dumas, agées de 24 à 28 ans en 1825. Ce trio de jeunes turcs a inséré sous la signature collective « R. » (pour « Rédaction »), en tête des dix-sept pages d'observations de Girou de Buzareingues, un « résumé spéculatif » de neuf pages ([Anonoyme], 1825), placé sous l'égide d'une référence aux recherches statistiques récentes du mathématicien Joseph Fourier (Fourier, 1). De même, le mémoire était-il ponctué d'un court *Nota bene* qui justifiait à nouveau sa publication et celle du « résumé spéculatif ». Il était encore prolongé par l'extrait d'un bulletin déjà paru, qui, cette fois, mettait en doute certaines des conclusions attribuées à Fourier (Bailly, 1824). La combinaison de ces moyens avait pour but rendre possible la publication du « résumé spéculatif ». On devine les jeunes rédacteurs frappés par la puissance d'une combinatoire qu'ils avaient extrapolé des observations de Girou de Buzareingues. Il est vrai que leur commentaire détaillait ce qui rétrospectivement aujourd'hui peut apparaître comme un prototype de la répartition d'une variable nominale selon deux variables centrées analogues. Il s'agissait en tout cas de mettre en évidence la variabilité de la génération selon une combinaison de deux mêmes mesures appliquées au père et à la mère. On sait que ces différents éléments seront présents au moment de la mise au point de la régression linéaire par Francis Galton, soixante ans plus tard, là encore à propos de la question de la transmission des qualités physiques dans le processus de la génération humaine (Galton, 1869 et 1886 ; Hald, 1998).

Girou de Buzareingues avait cru observer que les brebis d'âge « moyens » produisaient un agnelage équilibré entre les deux sexes, et que les brebis plus âgées ou plus jeunes produisaient plus de mâles. Les rédacteurs généralisèrent le constat selon une formule schématique de cette sorte :

- Mère *âgée*, sexe masculin fréquent.
- Mère d'*âge moyen*, rapports égaux.
- Mère *jeune*, sexe masculin fréquent.

Ils ajoutèrent cette extrapolation :

- Père *âgé*, sexe féminin fréquent.
- Père d'*âge moyen*, rapports égaux.
- Père *jeune*, sexe féminin fréquent.

Ils la combinèrent selon un principe d'addition des influences parentales. Nous résumons sous forme tabulaire *la liste ordonnée* des neufs cas possibles ainsi envisagés :

	Mère *jeune*	Mère d'*âge médian*	Mère *âgée*
Père *jeune*	Rapports égaux	S. *féminin* fréquent	Rapports égaux
Père d'*âge médian*	S. *masculin* fréquent	Rapports égaux	S. *masculin* fréquent
Père *âgé*	Rapports égaux	S. *féminin* fréquent	Rapports égaux

De même généralisèrent-ils l'autre critère agité par Girou de Buzareingues, l'état de santé des parents, en transcrivant le principe doxique qui veut que la force aille à la force :

- Mère *forte*, sexe féminin fréquent.
- Mère *moyenne*, rapports égaux.
- Mère *faible*, sexe masculin fréquent.
- Père *fort*, sexe masculin fréquent.
- Père *moyen*, rapports égaux.
- Père *faible*, sexe féminin fréquent.

De là, cette nouvelle combinaison :

	Mère *forte*	Mère *moyenne*	Mère *faible*
Père *fort*	Rapports égaux	S. *masculin* fréquent	S. *masculin* fréquent
Père *moyen*	S. *féminin* fréquent	Rapports égaux	S. *masculin* fréquent
Père *faible*	S. *féminin* fréquent	S. *féminin* fréquent	Rapports égaux

Après la publication de 1825 de l'article et de son faire valoir téméraire et collectif, Girou de Buzareingues n'a cessé d'accumuler les observations empiriques et les vues générales, les faisant valoir à l'Académie des sciences dont il était correspondant. Il a ainsi lu le début d'un *Mémoire sur la génération*, pendant la séance du 17 janvier 1825 (*PV*, t. VIII, p. 176). Au début de l'année 1827, il a adressé onze ouvrages à la Compagnie (*PV*, t. VIII, p. 483). Et au terme d'une longue série d'envois à la société savante, parfois de lectures en séance données par lui-même ou bien en son nom, il a communiqué à ses confrères, le 10 novembre 1828, un *Mémoire sur la distribution et les rapports des deux sexes dans le Royaume* où il analysait les chiffres d'état civil français au cours de la décennie précédente. Sa conclusion tenait en ce principe : « le sexe masculin est le résultat de la prédominance de la force motrice » (Girou de Buzareingues, 1828, p. 309). Un mois plus tard, le 15 décembre 1828, il soumettait son ouvrage de synthèse, intitulé *De la Génération*, au prix Montyon de physiologie expérimentale alors même qu'il se présentait à l'élection en vue de la place laissée vacante après la mort de Louis Bosc. Girou de Buzareingues ne fut pas élu, et le rapport de la commission du prix, lu le 8 juin 1829, lui attribua une mention honorable pour ses expériences, cela en dernière position après un prix et cinq autres mentions (*PV*, t. IX, p. 264-265).

Le couperet du calcul des probabilités

Entre temps le géomètre Poisson avait repris le flambeau du marquis de Laplace, mort le 5 mars 1827. Le 9 février 1829, en effet, devant une assemblée académique où, comme à l'accoutumé, ne figuraient pas moins de cinq des six membres de la Section de mathématiques et autant d'astronomes, où une douzaine des présents étaient liés au Bureau des Longitudes, mais où par extraordinaire le Secrétaire perpétuel Joseph Fourier était absent (*PV*, t. IX, p. 191-195), il donnait lecture d'un mémoire *Sur les proportions des naissances des garçons et des filles* dont la version publiée en 1830 rappelait la publication régulière, par ce Bureau, des chiffres de naissances enregistrées à l'état civil, selon le sexe et la légitimité.

Poisson continuait en déployant une longue démonstration resté célèbre dans l'histoire du calcul analytique des probabilités pour été un grand moment de mise point mathématique (Hald, 1998)

> « *Mais pour que les formules du calcul des probabilités dont il s'agit soient indépendantes de la loi de probabilité des écarts qui ne nous est pas donnée, il faut que les observations aient été faites en nombre considérable ; ce qui ne permet pas d'appliquer ces formules à la recherche du rapport des naissances annuelles des deux sexes, dont nous ne connaissons bien que les dix valeurs observées en France depuis 1817 jusqu'à 1826.* » (Poisson, 1830, p. 308).

Poisson renouait avec les premières recherches de Laplace sur la proportion des sexes à la naissance (Laplace, 1781). Pour des raisons à la fois de mathématique, de sources empiriques et de contexte historique, ses travaux avaient rapidement dérivé pendant les années 1780 vers d'autres objets et, en premier lieu, l'estimation du rapport entre le nombre des naissances et la population totale du royaume (Brian, 1994, p. 262-271). En 1830, l'essor de la circulation des chiffres et la disparition du fondateur de la théorie analytique du calcul des probabilités appelaient aux yeux de l'un de ses principaux successeurs un rappel à l'ordre exemplaire. Ce fut l'occasion de clarifier la question de l'estimation du paramètre d'une loi binomiale. Poisson constatait que les observations d'état civil pendant une dizaine d'années conduisaient à une estimation précise. Il indiquait la manière d'estimer par le calcul l'incertitude à laquelle il fallait se soumettre (A. Hald, 1998, p. 230-242). Poisson achevait son mémoire par une observation qui mérite l'attention des historiens des relations entre mathématiques et sciences sociales :

> « *Nous pouvons donc conclure qu'à l'époque actuelle et pour la France entière, la probabilité d'une naissance masculine n'éprouve que de très petites variations d'une année à une autre, et prendre pour sa valeur, la moyenne des dix années que nous avons considérées, c'est-à-dire, 0,5159. Dans l'ignorance où nous sommes de la cause qui rend prépondérantes les naissances des garçons, ce sera l'expérience seule qui pourra décider si cette probabilité variera d'avantage par la suite, où si elle demeurera à peu près constante. L'observation ne nous a pas encore appris si elle change dans une même année avec les saisons ; nous ne savons pas non plus si elle est la même chez les différentes nations ; nous savons seulement qu'elle dépend de l'état de la société, puisque le nombre des naissances hors de mariage influe sensiblement sur la proportion des naissances masculines et féminines.* » (Poisson, 1830, p. 307)

Probable réponse au géomètre avant la proclamation annuelle des prix, le 4 mai 1829, l'agronome aveyronnais était revenu à la charge, soumettant au jugement de la Compagnie un nouveau *Mémoire sur la distribution des mariages, des naissances et des sexes dans les divers mois*. Il l'annonçait comme une confirmation de ses expériences sur la génération, et il espérait le voir soumis à l'examen d'une nouvelle commission académique[5]. Le mathématicien Joseph Fourier et le physiologiste François Magendie furent désignés commissaires (*PV*, t. IX, p. 241). Mais le premier mourut le 16 mai 1830. Plus d'une année plus tard, le 1er août 1831, Girou de Buzareingues s'est rappelé au bon souvenir des académiciens en demandant un nouveau commissaire. Finalement, le 19 septembre 1831,

5. Archives de l'Académie des sciences, Pochette de séance du 4 mai 1829 (lettre de Girou fils datée du 4 mai 1829).

Girard, Damoiseau et Mathieu lurent un rapport très descriptif. Ils concluèrent en invitant l'agronome « à continuer ses intéressantes recherches » (*PV*, t. IX, p. 693-694). Dès le 22 août, un mois plus tôt, Girou de Buzareingues avait lu un *Mémoire sur le rapport des sexes dans le règne végétal.* Et, toujours le 19 septembre, un autre rapport académique sur cet ouvrage manifestait une certaine réserve (*PV*, t. IX, p. 685 et 697-698). Le compte de Girou de Buzareingues paraît avoir été soldé par les académiciens ce jour-là. L'agronome pourtant n'en restera moins jusqu'à sa mort un correspondant zélé.

Une expérience de physiologie à Tübingen vers 1827.

Il suffisait de traverser le Rhin pour voir que d'autres relations entre les mêmes sciences produisaient ailleurs qu'à Paris, et avec les mêmes ingrédients, des résultats à peine différents mais suffisamment affranchis des tensions académiques entre mathématiques et dénombrements empiriques. Il s'agit d'une thèse de médecine de l'Université de Tübingen intitulée *De qualitatibus parentum….* Elle a été soutenue le 1er septembre 1827 par Friedrich Notter (1801-1884) qui n'est pas connu aujourd'hui pour d'autres travaux scientifiques, mais pour son œuvre littéraire de traducteur de Dante et Cervantès, de journaliste et d'homme politique vers 1848 (Notter, 1827). On trouve de même une publication en langue allemande datée de 1828 dont le titre est à peu près la traduction du précédent, *Ueber die Eigenschaften…* (Notter, 1828). Cette fois, l'auteur sur la page titre en est le professeur de médecine vétérinaire et physiologiste Johann Daniel Hofacker (1788-1828). Celui-ci fut le « président » -- on dirait plutôt aujourd'hui le directeur -- de la thèse soutenue quelques mois plus tôt par Notter. Le nouveau docteur, pour sa part, apparaissait maintenant comme un collaborateur à l'ouvrage paru en allemand. Les deux versions, latine et allemande, sont très semblables. La préface de Hofacker à la seconde est curieusement datée du 10 septembre 1827. C'est dire que la publication en langue allemande, auprès d'un plus large public, et avec la mention explicite dans le titre de la portée des conclusions sur l'espèce humaine, fut décidée au plus tard dans les jours qui ont suivi la soutenance initiale. S'agirait-il d'un pillage ? Le critère paraît anachronique : l'usage voulait encore que les impétrants aient à défendre les thèses de leur mentor devant l'Université.

Quoiqu'il en ait été, il faut observer que Johan Daniel Hofacker était très attentif à ce qui se faisait alors à Paris. Il fut ainsi le traducteur du *Précis élémentaire de physiologie* de Magendie paru en allemand en 1826. Lui-même et les étudiants dont il a présidé les thèses ont étudié l'anatomie et la physiologie des animaux domestiques et de l'espèce humaine. Ce programme de travail collectif, apparemment interrompu par la mort prématurée de son animateur en 1828, faisait étonnamment écho, jusque dans le titre de la thèse de Notter, aux derniers paragraphes de Condorcet dans la Xème Epoque de l'*Esquisse*. Or, c'est précisément à Tübingen qu'avait paru sa première édition en langue allemande, cela dans le contexte assez connu aujourd'hui de l'intérêt des jeunes universitaires souabes pour les Lumières françaises et pour la Révolution (Condorcet, 1796). Les deux versions de la thèse s'appuient fermement sur Buffon, et citent explicitement le tome 22 de son *Histoire naturelle* commenté plus haut, où l'on retrouve en tête de ce volume, précisément, l'*Eloge de Buffon* où perçait dès 1788 le programme indiqué dans l'*Esquisse* et connu en langue allemande en 1796.

La thèse de Notter, comme sa version allemande, outre le nécessaire tribu aux autorités académiques et une citation de Shakespeare, repart de l'article de Girou de Buzareingues paru en 1825. Si elle indique précisément que Morel-Vindé avait le premier mis à l'épreuve le critère de l'âge de la mère, elle n'accorde pas d'attention à la différence entre le mémoire initial et la réélaboration des rédacteurs des *Annales des sciences naturelles*. Elle attribue ainsi à Girou de Buzareingues les spéculations de ses jeunes commentateurs. Mieux encore, c'est un véritable plan d'expérience que Notter a tiré de cette lecture et qu'il a appliqué à la lettre au dépouillement des registres de familles de la ville de Tübingen[6]. L'enquête a porté sur près de 2000 cas – l'effort était alors louable – et les résultats furent présentés avec prudence. Pourtant, ils tombent sous la critique mathématique que Poisson formulera en 1829 et qu'il publiera en 1830.

Itinéraire d'une affabulation.

Hofacker ne s'est pas contenté de proposer au principal éditeur universitaire de Tübingen de publier une traduction en langue allemande sous son nom. Dès avant la soutenance, il a aussi voulu alerter l'Allemagne savante. C'est pourquoi il a écrit à Johann Nepomuck Ehrhart von Ehrhartstein, le rédacteur du *Medicinisch-chirurgische Zeitung*, à Innsbruck. Dans les toutes dernières pages du supplément au dernier numéro de l'année 1827, Ehrhart a publié la communication de son confrère de Tübingen. En contradiction avec les passages de la thèse de Notter où les résultats de Girou de Buzareingues étaient comparés au cas par cas à ceux obtenus d'après les registres de Tübingen, Hofacker faisait sienne la découverte.

> « *Je saisis cette occasion en même temps pour vous proposer quelques notes à propos d'une dissertation publiée prochainement sous ma présidence en espérant qu'elles pourraient convenir en vue d'une annonce dans votre journal. Dans ce mémoire qui traite des qualités que les parents transmettent aux enfants, j'en suis venu naturellement à la détermination du sexe par différents éléments. Pour diverses raisons, je pensais depuis longtemps déjà que l'âge était à cet égard un élément important. J'ai extrait des registres de baptêmes, pour 2 000 enfants, les indications relatives à leur sexe et à l'âge du père et de la mère, et j'ai trouvé les principaux résultats suivants. [suivent 15 constats chiffrés]. Toutes ces observations et ces calculs sont établis avec la plus grande précision. Dans un souci de confirmation, j'ai laissé un bon ami procéder de même sur un village voisin, ce qui a livré des résultats tout à fait semblables. Manifestement, il s'agit là de lois tout à fait nouvelles, vers la découverte desquels j'ai été conduit par le hasard et quelque persévérance dans le calcul.* » (Hofacker, 1827, p. 398-399 ; notre traduction).

Un an et demi plus tard, c'est-à-dire après la lecture du mémoire de Poisson à l'Académie des sciences, la nouvelle imprimée en Autriche avait atteint Paris. Charles-Chrétien-Henri Marc, l'un des rédacteurs des *Annales d'hygiène publique et de médecine légale*, périodique qui lui aussi tentait de trouver une place pour la science expérimentale hors des mises en garde des mathématiciens les plus stricts de l'Académie des sciences (Lécuyer, 1977), répercuta la correspondance tendancieuse dans les colonnes de cette toute nouvelle revue. La note en français voulait traduire l'essentiel de la lettre allemande, mais faute d'en saisir la structure tabulaire sous-jacente seuls quatorze résultats y furent mentionnés : deux lignes de la version autrichienne furent ainsi confondues. On y comprend à peine quelles furent les classes employées pour les âges des parents (Marc, 1829).

La série de chiffres est imprimée pages 557-558 des *Annales d'hygiène*. Adolphe Quetelet quelques années plus tard est parti de cette feuille parisienne, laissant courir dans les diverses

---

6. La thèse présente des résultats de dépouillements effectués sur des registres des haras d'Etat de Marbach (avec l'autorisation du gouvernement). Toutefois, malgré les deux titres de l'ouvrage, ces compilations dans le cas des chevaux ne se présentent pas sous une forme aussi systématique que dans celui des hommes.

éditions de sa *Physique sociale* (1835, 1869) une pagination fausse (« p. 537 »). Elle a dérouté ses successeurs… et même ses éditeurs récents (1997). Dans ces conditions le tableau alors publié comporte ainsi non seulement une série d'erreurs de transcription et une lourde faute de traduction qui le rend presque incompréhensible, mais encore un vice profond : l'escamotage d'étape en étape du travail de construction empirique mené très attentivement depuis Morel-Vindé (1814) et jusqu'à Notter (1827).

Ce n'est pas tout, Adolphe Quetelet, en 1835, a dans les pages de sa *Physique sociale* noué pour la première fois les noms de Sadler et de Hofacker en une même formule – « les recherches de MM. Hofacker et Sadler » – accordant au premier, en 1829, la primeur d'une observation dont il ne savait comment elle avait été établie et au second, en 1830, le mérite d'avoir mobilisé des chiffres assez probants pour appuyer la conjecture de son prédécesseur. On scrutera en vain dès lors pendant presque un siècle la répartition des sexes à la naissance pour y chercher la marque d'une loi de Hofacker-Sadler qui lierait le taux de masculinité des nouveau-nés et l'écart d'âge de leurs parents (Brian et Jaisson, 2007a et 2007b).

Grand organisateur de l'observation statistique administrative en Belgique puis en Europe, mais aussi grand prédateur de chiffres, Quetelet se révèle en 1835 tout à fait désinvolte à l'égard des procédés de compilation et de calcul. L'apparition de la lettre de Hofacker à Paris et sa combinaison forcé avec chiffres de Sadler paraissent avoir permis au statisticien belge de contourner la sentence de Poisson[7]. Le succès de la fable de deux découvertes parallèles hors de la scène française a eu pour effet de favoriser une réouverture de la question empirique soulevée par Condorcet, et semble-t-il favorisée par Fourier. Au passage, réduite à une conjecture statistique issue de l'Angleterre anglicane réformatrice et de l'Allemagne romantique, elle perdait le sens d'un horizon des progrès de l'esprit humain affranchi des causes de la superstition.

Le nom de Hofacker, porté depuis lors par la tradition statistique appelle une réflexion supplémentaire. En effet, le contexte intellectuel à Tübingen vers 1830 n'est plus celui de disons de la jeunesse de Hegel aux temps de la traduction précoce de l'*Esquisse* de Condorcet. Au temps du succès de la *Physique sociale*, le nom de Hofacker n'était pas connu en Europe comme celui d'un physiologiste mais comme celui d'un mystique homonyme Ludwig Hofacker (1798-1828) qui prêchait un renouveau évangélique fondé sur l'idée que, Satan étant « le suprême *Aufklärer* », les Lumières devait annoncer la fin des temps. Dès les années 1830 (soit quelques années après la mort en 1828 des deux homonymes), le théologien anti-rationaliste fut certainement plus célèbre parmi les élites de langue allemande que le physiologiste empiriste vite oublié. Ce constat laisse perplexe quand on songe au relatif succès du syntagme « Hofacker-Sadler » après 1850.

Effet d'amnésie induit par la circulation internationale des dénombrements, enthousiasme de statisticien conquérant désireux de couper court, ou bien hypocrisie positiviste bien entendue, quoiqu'il en soit trois aspects fondamentaux de la question de la régularité du rapport des sexes à la naissance seront désormais escamotés et transformés en un poncif statistique répéré par l'expression « Hofacker-Sadler » : (1) quelle philosophie la possibilité du calcul présuppose-t-elle (pessimiste, optimiste ou mélioriste ?) ; (2) sur quelles traditions intellectuelles le calcul est-il fondé ? (théologie, calcul des probabilités ou bien science sociale ?) ; (3) quels présupposés le calcul recouvre-t-il en ce qui concerne l'animalité des hommes et la distinction entre animaux domestiques et animaux en général ?

---

7. Quetelet a mentionné en passant le mémoire de Poisson publié en 1830 (p. 62 de l'éd. de 1997), et la communication à l'Académie des sciences de Paris d'un certain Giron de Buzareignes (*sic*). Il s'agit bien sûr de Girou de Buzareingues et, tout porte à le croire, de sa lecture interrompue en 1825 (p. 63-64 de l'éd. de 1997).

Un autre itinéraire non moins périlleux

Un trait caractérise les recherches de Darwin, et celles développées dans son sillage, si on les compare aux travaux continentaux. Elles sont conformes à la résolution adoptée en 1839 par la *British Association for the Advancement of Science* : l'espèce humaine y est tenue pour un objet d'histoire naturelle (Browne, 1995, p. 421 ; [BA], 1840). Condorcet n'en aurait pas moins attendu des sciences pendant la X$^{\text{ème}}$ époque utopique qu'il avait assigné à l'esprit humain (Condorcet, 1795 [2004]). Mais l'ébauche de sa perspective mélioriste avait suscité, on s'en souvient, la réaction de Malthus qui, à l'encontre d'une spéculation si heureuse sur l'avenir, avait agité l'incohérence entre la progression géométrique de la population et celle arithmétique des subsistances (Malthus, 1798). Or c'est précisément en étendant le constat du Révérend que Darwin a défini le concept de sélection naturelle (Darwin, 1859 ; Tort, 1996). Ainsi, en trois temps historiques et à la manière d'une ruse de la raison, l'horizon empirique darwinien s'est-il formé. Dans un premier lieu, à Paris à la fin du XVIIIe siècle, dans le contexte d'une mise en suspens métaphysique de la théologie naturelle à la Süssmilch et à l'encontre de la trop grande généralité des conclusions du calcul analytique des probabilités du jeune Laplace, Condorcet a indiqué les conditions morales d'une physique générale de l'homme et de l'animal pour illustrer sa conception des progrès asymptotiques de l'esprit humain : le sujet de l'histoire – l'esprit humain – étendrait ainsi son horizon. Dans un second temps, à Londres autour de 1800, la réception de l'*Esquisse d'un tableau des progrès de l'Esprit humain* a suscité chez Malthus une réaction néo-providentialiste qui a consisté à dénier à l'esprit humain ce statut de sujet de l'histoire et à lui opposer la problématique de la population et des subsistances issue de l'économie politique du XVIIIe siècle et propre, dans ce contexte, à l'espèce humaine. Plus d'un demi-siècle plus tard, conformément à une résolution de la société savante britannique qui consolidait l'horizon empirique attendu par Condorcet, Darwin a élaboré le concept de sélection naturelle en étendant à toutes les espèces la contradiction qu'avait agitée Malthus précisément à l'encontre de la philosophie de l'histoire que présupposait l'action même de la Société pour l'avancement des sciences.

On sait que Charles Darwin fréquentait les publications de Quetelet, mais qu'il se défiait de sa méthode. On sait aussi que Darwin mobilisait un matériel empirique considérable par des canaux les plus divers (Browne, 2002) : observations personnelles, commerce savant de longue haleine avec des interlocuteurs privilégiés, informations transmises par des correspondants de loin en loin, ou même glanes occasionnelles et interactions opportunes quand, par exemple, son traducteur en langue néerlandaise, Hermanus Hartogh Heys van Zouteveen, lui a procuré entre les deux premières éditions de *The Descent of Man* (1871 et 1874) certains chiffres sur la proportion des sexe parmi les naissances des blancs et des esclaves du Cap de Bonne Espérance tirés de Quetelet, non sans réserve, et dont les éditeurs de la *Physique sociale*, il y a encore quelques années, ont été bien en peine d'identifier la source (Quetelet, 1997, p. 60-61). Darwin comme nombre de savants contemporains brassait des chiffres sans souci critique ni grande attention à leur origine.

De quoi disposait Darwin à propos de la proportion des sexes à la naissance ? La comparaison avec deux aînés, Quetelet et le médecin hygiéniste et statisticien français Louis-René Villermé est éclairante : l'un et l'autre ont commenté les mêmes ouvrages publiés en allemand, en anglais ou en français (Quetelet, 1835 et 1869 ; Villermé, 1932). Laplace, Poisson, Hofaker (c'est-à-dire Notter) et Girou de Buzareingues, par exemple, apparaissent ici comme là. Seuls les canaux d'information diffèrent et par voie de conséquence la place qu'occupent les mentions à ces auteurs dans l'économie générale des ouvrages de ceux qui y font référence. A cet égard, un ouvrage particulier paraît avoir guidé l'approche darwinienne de la répartition des sexes à la naissance. Il s'agit d'un mémoire paru plus de trente ans auparavant, et adressé par Charles Babbage au rédacteur de l'*Edinburgh Journal of Science* en 1829 (Babbage, 1829). Or on le sait, Babbage fut l'un des principaux interlocuteurs de Darwin à propos des questions de statistique des naissances et des travaux de Quetelet (Browne, 1995, p. 385).

Le mathématicien Babbage, dont on retient le plus souvent aujourd'hui la contribution à l'histoire de la mécanisation du calcul, était un personnage éminent de l'entourage des parrains scientifiques et intellectuels de Darwin (Browne, 1995). C'est ici le monde de l'*Analytical Society*, fondée en 1812 à Cambridge, qui a assuré la promotion de la conception analytique des mathématiques en vigueur sur le continent, alors que dominaient en Grande-Bretagne le modèle synthétique des sciences mathématiques newtoniennes. Babbage et ses proches – John F. W. Herschel, George Peacock, puis plus tard Augustus de Morgan et George Boole – étaient à ce titre des plus attentifs à toutes les innovations françaises dans les sciences mathématiques de la fin du XVIIIe et du début du XIXe siècle. Réformer ainsi la science anglaise au moyen de la science continentale, dans le contexte militaire et politique de cette période, ce n'était pas seulement adopter des résultats et des méthodes trouvés ailleurs: c'était aussi sortir la science anglaise du repli induit par la période du blocus napoléonien et rechercher délibérément en France et en Allemagne tout ce qui pouvait contribuer à renforcer l'affirmation d'une jeune génération d'hommes de science. Le geste le plus manifeste de la Society fut la traduction d'un des meilleurs manuels d'analyse écrit en français à cette époque, celui de Silvestre Lacroix qui avait longtemps secondé Condorcet dans l'enseignement (Lacroix, 1802 ; 1816) : il s'agissait de promouvoir les notations leibnizienne du calcul intégral, plus puissantes il est vrai, au détriment des notations newtoniennes préservées en Angleterre (Babbage, 1864 ; Enros, 1979 ; Durand-Richard, 1996 et 2001).

L'article de 1829 provient d'une tension entre Babbage et De Morgan à propos de ce que les mathématiciens auraient à dire des assurances. Les travaux de Condorcet et de Laplace, cinquante ans plus tôt à Paris, avaient eux-mêmes touché cette question. L'un et l'autre, alors qu'ils oeuvraient à l'élaboration du calcul analytique des probabilités, remplissaient en effet de loin en loin à l'Académie des sciences des fonctions d'expertise car cette Compagnie savante, à la fin de l'Ancien Régime, avait à se prononcer sur le bien fondé de divers projets de sociétés d'assurances (Condorcet, 1994 ; Brian, 1994). A un demi-siècle de distance, ce sont des circonstances homologues ici et là. Elles sont toutefois inscrites dans des cadres institutionnels et des divisions sociales différentes des tâches de contrôle sur le commerce des assurances (Campbell-Kelly, 1994). De Morgan avait critiqué le recueil que Babbage, alors au service de la Protector Life Assurance Society, avait consacré à la comparaison des différentes formules d'assurance sur la vie (Babbage, 1826 ; voir aussi De Morgan, 1838). L'article paru à Edinburgh en 1829 se présente sous la forme d'une lettre à l'éditeur du Journal accompagnée d'un dossier de tableaux statistiques rassemblés par l'auteur auprès de quelques correspondants informés. En 1826, Babbage avait été l'avocat du développement de l'industrie assurantielle. Il avait plaidé à ce sujet la confiance dans l'application du calcul des probabilités. Trois ans plus tard, il ironisait sur les réserves de De Morgan exprimées entre-temps. Son critique ayant rappelé la variabilité des observations publiées par le Bureau des longitudes à Paris, Babbage entendit montrer qu'il n'ignorait pas les variations observées quant à la proportion des deux sexes à la naissance, ni leur importance pour le commerce des assurances tant « facts and accurate enumerations are the great and only bases on which such transaction can securely rest » (Babbage, 1829). Ce faisant, Babbage a rappelé, mentionnant Laplace, que la proportion des sexes parmi les enfants de l'hospice de Paris était exceptionnelle. Dans ces conditions, il a discuté, chiffres à l'appui, trois cas de figures connexe : celui des naissances illégitimes, celui des différences entre la ville et la campagne et celui des enfants morts-nés. Le tableau de l'objection que Laplace s'était faite à lui-même presque vingt ans plus tôt, celui que Fourrier avait réélaboré en appelant de ses vœux toujours plus d'observations organisées, a trouvé sous la plume de Babbage des éléments complémentaires. Le phénomène, il en convenait, présentait des indices de variabilité. Il laissait entendre qu'il y avait là, non pas des raisons de douter mais autant d'examens empiriques à entreprendre. Babbage, ensuite, livrait à l'attention des savants les tableaux que lui avaient procuré trois de ses informateurs depuis 1826, l'un compilateur des statistiques du Foundling Hospital de Dublin, l'autre, 'Chef de Division et Directeur de Bureau Statistique dans la ministére de l'Interieure' (sic [en français sous la plume de B.]) de Westphalie, et le troisième, chef du bureau de statistique à Berlin : Johann Gottfried Hoffmann. Il y ajoutait la mention d'une recherche présentée en 1827 à l'Académie des sciences de Paris sans en indiquer l'auteur. Ces éléments allemands et français provenaient de son récent voyage en Europe.

Or Johann Gottfried Hoffmann était très préoccupé par la statistique des confessions en Prusse et tout particulièrement par la mesure de ce qu'il qualifiera à la fin de sa vie de « Judenfrage », forgeant ainsi la matière empirique d'une réflexion appelée à prendre une terrible ampleur sur la place des juifs dans la société prussienne (Hoffmann, 1842 ; 1844). C'était un économiste de première importance en son temps (Schuster, 1908). Ses publications statistiques contribuèrent fortement à la légitimation de l'anti-sémitisme lettré en Allemagne (Keval, 1999). Ses articles des années 1840 montrent que dès les années 1820, il établissait une statistique des juifs de Prusse. Il avait communiqué des éléments préliminaires de ces mêmes travaux à Babbage qui les a publié en annexe de sa lettre au Edinburgh Journal of Science. Babbage les a introduit en ces termes après la discussion des variations de la proportion des naissances selon les sexes déjà indiquées avant lui :

> "*I shall notice one other circumstance connected with this subject. It is the remarkable excess of males amongst the children of the Jews of Prussia. For every ten thousand females born amongst them there are 11 292 males [53,03%].*" (Babbage, 1829, p. 91)

Quel est l'implicite de cette évocation d'"autres circonstances" ? Le contexte demeure ici d'ordre confessionnel et pas encore tout à fait racial au sens d'une conception naturaliste des humains. Parce que l'arrière plan de la discussion de la proportion des sexes à la naissance demeure sous la plume de Babbage la critique de la physico-théologie si puissante en Angleterre au XVIIIe siècle, la question de la polygamie était en effet ici en jeu. C'est pourquoi Babbage prolongeait immédiatement son argument en amalgamant les religions juive et musulmane :

> "*It would be interesting to examine this fact amongst the Jews of other countries, and still more so, could we procure any correct enumeration of births in any country in which the Mahometan religion prevails.*" (Babbage, 1829, p. 91)

Darwin, en 1871, quoiqu'il s'appuiera fortement sur le texte de Babbage, ne retiendra pas cette formulation. Mais il y reviendra d'une autre manière (Brian et Jaisson, 2007b). Babbage, toujours en 1829, est ensuite passé à une autre observation curieuse, allant sans transition des religions compatibles avec la polygamie aux animaux domestiques:

> "*I cannot conclude this subject, without recalling to your notice a statement, in the History of the Academy of Sciences of Paris for the year 1827. It is stated as the result of some experiments lately tried, that in a flock of sheep consisting of 71 females and 61*

males, by selecting strong females and young males, and by feeding the females high and not the males, the result was amongst the births

	Males	Females	
	53	84	
by the reverse process	80	50”	(Babbage, 1829, p. 91)

On aurait eu peine, en 1829, à savoir de quoi Babbage parlait ! En effet, le compte rendu d’activité pour l’année 1827, traditionnellement appelé « Histoire de l’académie des sciences » dans l’environnement de l’Académie parisienne, rédigé par le secrétaire perpétuel pour les sciences physiques, Georges Cuvier, ne paraîtra qu’en… 1831 (Cuvier, 1831). Babbage évoque donc, en 1829, une séance dont on lui aura parlé lors de son séjour à Paris, un an plus tôt, ou bien il reprend des notes prises à la lecture d’un mémoire qu’on lui aurait procuré (rien n’indique à notre connaissance qu’il ait assisté aux lectures données à l’Académie de Paris). Bref, il fait montre de compétence en réponse à De Morgan, bien qu’en fait sa relation soit quelque peu fautive : les chiffres diffèrent du compte-rendu paru deux ans plus tard ; de plus, le cas des « 71 femelles et 61 mâles » relève d’une troisième expérience, que nous qualifierions aujourd’hui de groupe témoin. Mais ce n’est pas tout, le volume académique de 1831 comme les procès verbaux des réunions parisiennes de l’année 1827 prouvent sans ambiguïté que les expériences évoquées par Babbage sont celles de Girou de Buzareingues.

> « *Des expériences curieuses, non seulement pour l’agriculture, mais pour la physiologie générale, sont celles de M. Girou de Busareingues (sic), sur la procréation des sexes. C’est du plus ou moins de vigueur comparative des individus que l’on accouple, que dépend selon lui le sexe du produit. Si l’on veut avoir plus de femelles, il faut employer des mâles jeunes et des femelles dans l’âge de la force, et nourrir celles-ci plus abondamment que ceux-là. Il faut faire l’inverse si l’on veut produire plus de mâles. Avec le premier procédé l’on a obtenu d’un agnelage 84 femelles contre 53 mâles ; et avec le second, l’on a eu 55 brebis contre 80 mâles ; tandis qu’une égalité de force et de nourriture avait donné dans le même troupeau 71 femelles et 61 mâles. Les oiseaux suivent la même loi que les moutons. Dans la même basse-cour, les plus fortes femelles procurent un nombre d’individus de leur sexe plus grand que les petites ; les jeunes femelles qui n’ont pas acquis un développement précoce, donnent plus de mâles.* » (Cuvier, 1831, p. CLXXXVII)

L’article de Babbage paru en 1829 nous offre un témoignage des bases de la discussion entre ce mathématicien et son cadet naturaliste, Darwin, à propos de la répartition par sexe des naissances. On peut en trouver aisément de multiples indices dans *The Descent of Man* (1871) en repérant les références faites aux auteurs et aux questions indiquées par Babbage lui-même (Brian et Jaisson, 2007b). Les deux auteurs ont adopté la position promue dans la résolution prise par la British Association en 1839 à propos de la naturalité de l’homme et elle présuppose une syntaxe particulière des choses connues : la proportion des sexes à la naissance offrirait à l’échelle de l’humanité une constance très générale ; la question de la variabilité empirique à l’intérieur de l’espèce humaine resterait à traiter ; les variations établies toucheraient en général les naissances naturelles et les enfants morts-nés ; un continuum de cas se dessinerait qui s’ordonnerait depuis les peuples monogames jusqu’aux diverses espèces animales, en passant par les peuples polygames et les espèces domestiquées. C’est par cette voie que la statistique de la variabilité humaine a conduit des recherches de la première moitié du XIXe siècle à Charles Darwin, puis à Francis Galton et, pour ce qu nous intéresse, à Ronald. A Fisher.

L'expérience statistique pré-quetelésienne

A l'issue d'une enquête récemment menée sur trois siècles de calculs de la proportion des sexes à la naissance, nous avons qualifié de régime *pré-quetelésien* le début de la période étudiée dans cet article, quand parmi les conceptions statistiques ne primaient ni la référence à une valeur centrale, ni l'analogie entre la structure de la variabilité du phénomène et celle des erreurs que sa mesure comporte (Brian et Jaisson, 2007b). Lire Condorcet, Laplace ou Babbage sans les considérer comme des savants d'une telle époque expose l'historien à des anachronismes périlleux. Restituer cette caractéristique, c'est se donner le moyen de saisir la finesse des constructions de tant d'autres savants qui leurs étaient contemporains. Le régime qu'on peut ensuite qualifier de *quetelésien* est mieux connu : l'activité des savants y est gouvernée pas la normalisation des observations selon la théorie de la moyenne et des erreurs formulée par Quetelet et portée par l'essor des institutions statistiques européennes. Ce régime s'instaure pendant la seconde moitié du XIXe siècle : plusieurs auteurs ont déjà observé que les années 1840 marquaient à cet égard un tournant (Daston, 1986 ; Porter, 1986 ; Hacking, 1990).

Bien avant que les statisticiens, approfondissant la voie ouverte par Darwin et Galton, ne se dotent au XXe siècle, pendant l'entre-deux-guerres, d'un attirail de tests à appliquer aux mesures de la variabilité des phénomènes, des savants agronomes, des physiologistes et des médecins ont cherché à explorer par des observations systématiques, organisées et comparées à traiter des questions qui touchaient à la génération et prolongeaient en cela les maigres indications laissées par Condorcet ou Fourier. Mais faute d'une « langue de la science » qui aurait fait s'entendre un peu mieux les savants européens du XIXe siècle, ces recherches ne se sont pas consolidées les unes les autres. Au bilan, leur circulation apparaît même assez pitoyable. Il a donc fallu que le simplisme quetelésien s'imposât aux statisticiens, et que le souci darwinien – antagoniste – de traquer la variabilité des phénomènes devînt la marque de l'école britannique de statistique pour se soit formé un tel idiome. Si bien que les tentatives des statisticiens expérimentalistes du début du XIXe siècle peuvent après coup être qualifiés de *plans d'expérience pré-quetelésien*.

Bibliografía

[ANONYME] « Résumé spéculatif », Annales des sciences naturelles, t. V, 1825, p. 21-29 (accompagné par un « Nota bene », p. 46).

[BA] Report of the Ninth Meeting of the British Association for the Advancement of Science held at Birmingham in August 1839, London, 1840.

BABBAGE, CHARLES, "A Letter to the Right Hon. T. P. Courtenay, on the proportionate number of Births of the two Sexes under different Circumstances", Edinburgh Journal of Science, vol. 1, 1829, p. 85-104 (art. XIII ; daté du 7 mai 1829).

BABBAGE, CHARLES, A Comparative View of the various Institutions for the Assurance of Lives, London, Mawman, 1826.

BABBAGE, CHARLES, Passages from the Life of a Philosopher, London, Longman, 1864 (éd. récente: New Brunswick, Rutgers University Press, 1994).

[BABBAGE, CHARLES], The Works of Charles Babbage, New York, New York University Press, 1989, 11 vols. (ed. by Martin Cambell-Kelly).

BAILLY (Dr.), « Considérations sur l'influence des circonstances extérieures dans les conceptions et les naissances masculines et féminines », Annales des sciences naturelles, t. V, 1825, p. 47-49 (repris du Bulletin de la Société de Philomatique, octobre 1824).

BRIAN, ÉRIC , La Mesure de l'État. Administrateurs et géomètres au XVIIIe siècle, Paris, Albin Michel, 1994 (Staatsvermessungen. Condorcet, Laplace, Turgot und das Denken der Verwaltung, Wien, Springer Verlag, 2001).

BRIAN, ÉRIC ; JAISSON, MARIE, Le Sexisme de la première heure. Hasard et sociologie, Paris, Raisons d'agir, 2007a.

BRIAN, ÉRIC ; JAISSON, MARIE, The Descent of Human Sex Ratio at Birth. A Dialogue between Mathematics, Biology and Sociology, Dirdrecht, Springer, 2007b.

BROWNE, JANET, Charles Darwin. Voyaging, vol. 1, London, Jonathan Cape, 1995 ; Charles Darwin. The Power of Place, vol. 2, London, Jonathan Cape, 2002.

CAMPBELL-KELLY, MARTIN, "Charles Babbage and the assurance of lives", Annals of the history of computing, vol. 16, n°3, 1994, p. 5-14.

CONDORCET, « Eloge de Buffon », Histoire de l'Académie royale des sciences, année 1788, Paris, 1791, p. 50-84 (cet éloge figure en tête du tome 22 de l'édition de l'Histoire naturelle de Buffon, éd. de l'an VIII, p. 5-64).

CONDORCET, Entwurf eines historischen Gemähldes der Forschritt des menschlichen Geistes, Tübingen, Cotta, 1796 (trad. de l'édition de l'Esquisse d'un tableau des progrès de l'esprit humain, Paris, An III, par Ernst Ludwig Posselt).

CONDORCET, « Mémoire sur le calcul des probabilités » (en 6 articles et 4 livraisons), Mémoires de l'Académie royale des sciences. Années 1781, 1782, 1783, 1784, Paris, 1784, p. 707-728 ; 1785, p. 694-691 ; 1786, p. 539-559 ; 1787, p. 454-477 (repris dans Condorcet, Arithmétique politique. Textes rares ou inédits, Paris, INED, 1994, p. 387-448).

CONDORCET, Tableau historique des progrès de l'esprit humain. Projets, Esquisse, Fragments et Notes (1772-1794), Paris, Ined, 2004.

CUVIER, GEORGES, « Analyse des travaux de l'Académie royale des sciences pendant l'année 1827. Partie physique », Mémoires de l'Académie des sciences de l'Institut de France, Paris, Firmin-Didot, 1831, tome 10, p. CIII-CXC.

DARWIN, CHARLES, The Descent of Man and Selection in Relation to Sex, London, John Murray, 1871 (2nd ed., 1874 ; édition de référence ultérieure, 1882 ; traduction néerlandaise historique par Hermanus Hartogh Heys van Zouteveen, De afstamming van den mensch, en de seksueele teeltkeus, Delft, Ykema & Van Gyn, 1871-1872 ; traduction française de réference, La Filiation de l'homme et la sélection liée au sexe, Paris, éditions Syllepse, 2000).

DARWIN, CHARLES, The variation of animals and plants under domestication, London, John Murray, 1868.

DASTON, LORRAINE, Classical probability in the Enlightenment, Princeton, Princeton University Press, 1988.

DE MORGAN, AUGUSTUS, An essay on probabilities, and on their application to life contingencies and insurance offices, London, Longman etc., 1838 (partial reprinted in Jenkins and Yoneyama, 2000, vol. 5).

DURAND-RICHARD, Marie-José, « L'École algébrique anglaise. Les conditions conceptuelles et institutionnelles d'un calcul symbolique comme fondement de la connaissance », dans Catherine Goldstein, Jeremy Gray and Jim Ritter (eds), L'Europe mathématique. Histoires, mythes, identités, Paris, éd. de la Maison des sciences de l'homme, 1996, p. 445-477.

DURAND-RICHARD, MARIE-JOSE, « Révolution industrielle. Logique et signification de l'opératoire", Revue de synthèse, n°2-3-4, 2001, p. 319-346.

DUVAL, JULES, Girou de Buzareingues, Paris, Hennuyer, 1858.

ENROS, PHILIP CHARLES, The Analytical Society : mathematics at Cambridge University in the early nineteenth century, Toronto, Univesity of Toronto (Ph.D.), 1979.

Fisher, Ronald A., Statistical methods for Research Workers, Edinburgh, Oliver and Boyd, 1925 (rééd. Statistical methods, experimental design, and scientific inference, Oxford/New York, Oxford University Press, 1990).

FISHER, RONALD A., "The arrangement of field experiments", Journal of the Ministry of Agriculture of Great Britain, 1926, vol. 33, p. 503-513.

FISHER, RONALD A. AND WISHART, JOHN, "The arrangement of field experiments and the statistical reduction of the result", Imperial Bureau of Soil Science, Technical communication n°10, London, HMSO, 1930a.

FISHER, RONALD A., The Genetical Theory of Natural Selection, Oxford, Clarendon Press, 1930b (2nd ed., New York, Dover, 1958).

FISHER, RONALD A., The design of experiments, Edinburgh, Oliver and Boyd, 1935.

FOURIER, JOSEPH, « Notions générales sur la population » dans Recherches statistiques sur la ville de Paris et le département de la Seine... Recueil de tableaux dressés et réunis d'après les ordres de M. le Comte de Chabrol, préfet du département, Paris, Ballard puis imprimerie royale, 4 volumes, 1821-1829, ici 1821, p. IX-XLVIII.

FOURIER, JOSEPH, « Mémoire sur la population de la Ville de Paris depuis la fin du XVIIe siècle », ibid., ici 1823, p. XIII-XXVIII.

FOURIER, JOSEPH, « Règle usuelle pour la recherche des résultats moyens d'un grand nombre d'observations », Bulletin des sciences mathématiques, physique et chimique, 1824, n°2, p. 88-90.

GALTON, FRANCIS, Hereditary genius : an inquiry into its laws and consequences, London, Macmillan, 1869.

GALTON, FRANCIS, "Regression towards mediocrity in hereditary stature", Journal of the Anthropological Institute, vol. 15, 1886, p. 246-263.

GIRARDIN, JEAN, Notices biographiques sur de Morel-Vindé, d'Arcet et Mathieu de Dombasle, Rouen, Péron, 1845.

GIROU DE BUZAREINGUES, CHARLES, « Observations sur les rapports de la mère et du père avec les produits, relativement au sexe et à la ressemblance », Annales des sciences naturelles, t. V, 1825, p. 30-46.

GIROU DE BUZAREINGUES, CHARLES, De la Génération. Suivi d'un Mémoire sur la distribution et le rapport des sexes en France, Paris, Huzard, 1828.

HACKING, IAN, The Taming of Chance, Cambridge, Cambridge, University Press, 1990.

HALD, ANDERS, A History of Mathematical Statistics from 1750 to 1930, New York, Wiley, 1998.

[HOFACKER, JOHAN DANIEL], « Aus einem Schreiben des hn. Professors Hofacker in Tübingen an den herausgeber des med. chir. Zeitung », Medicinisch-chirurgische Zeitung, 1827, supplément au n°100 (13 décembre 1827), p. 398-399.

HOFFMANN, JOHANN GOTTFRIED, „Übersicht der bei dem statistischen Bureau zu Berlin vorhandenen Nachrichten über die Anzahl und Vermehrung der Juden im preussischen Staate", Philologische und historische Abhandlungen der königlichen Akademie der Wissenschaften zu Berlin. Aus dem Jahr 1842, 1844, p. 224-243.

HOFFMANN, JOHANN GOTTFRIED, Zur Judenfrage: statistische Erörterung. Anzahl ... der Juden im Preussischem Staate nach einer Vergleichung der Zählungen im Jahren 1840 und 1822, Berlin, Lesekabinet, 1842 (29 p.).

JENKINS, DAVID AND YONEYAMA, TAKAU, The History of Insurance, London, Pickering & Chatto, 2000 (8 vols.).

KEVAL, SUSANNA, Die schwierige Erinnerung : deutsche Widerstandskämpfer über die Verfolgung und Vernichtung der Juden, Frankfurt/Main, Campus-Verlag, 1999.

LACROIX, SILVESTRE FRANÇOIS, An elementary treatise on the Differential and Integral Calculus…. Translated from the French, Cambridge, Deighton, 1816 (traduit de Traité Élémentaire de Calcul Différentiel et de Calcul Intégral. Précédé de réflexions sur la manière d'enseigner les Mathématiques, et d'apprécier dans les examens le savoir de ceux qui les ont étudiées, Paris, Duprat, 1802).

LAPLACE, PIERRE SIMON, « Mémoire sur les probabilités », Mémoires de l'Académie royale des sciences. Année 1778, Paris, impr. royale, 1781, p. 227-332 (Œuvres, t. IX, Paris, Gauthier-Villars, 1893, p. 383-485).

LAPLACE, PIERRE SIMON, Théorie analytique des probabilités. Paris, 1812 (éd. de référence : 1820 ou Œuvres, t. VII, Paris, Gauthier-Villars, 1886).

LAPLACE, PIERRE SIMON, Essai philosophique sur les probabilités, Paris, 1814 (éd. référence : 1825 ou Essai philosophique…, Paris, Gauthier-Villars, 1921, 2 vol.).

LECUYER, BERNARD-PIERRE, « Démographie, statistique et hygiène publique sous la monarchie censitaire », Annales de démographie historique, Paris, 1977, p. 215-245.

MALTHUS, THOMAS ROBERT, An essay on the principle of population : as it affects the future improvement of society with remarks on the speculations of Mr. Godwin, M. Condorcet, and other writers, London, J. Johnson, 1798 (trad. fr. Essai sur le principe de population en tant qu'il influe sur le progrès futur de la société, avec des remarques sur les théories de Mr Godwin, de M. Condorcet et d'autres auteurs, Paris, Ined, 1980).

MARC, CHARLES-CHRETIEN-HENRI, « Extrait d'une lettre du professeur Hofacker au rédacteur de la Gazette médico-chirurgicale d'Innsbruck (autrefois de Salzbourg) », Annales d'hygiène publique et de médecine légale, 1829, t. Ier, 2ème partie, p. 557-558.

MOREL-VINDE, CHARLES-GILBERT TERRAY DE, Observations sur la monte et l'agnelage, Paris, Huzard, 1813 ; Suite des Observations sur la monte et l'agnelage, Paris, Huzard, 1814 ; Seconde suite des Observations sur la monte et l'agnelage, Paris, Huzard, 1815 ; Troisième et dernière suite des Observations sur la monte et sur l'agnelage, suivie de quelques considérations générales sur les encouragemens [sic] à donner à l'agriculture, Paris, Huzard, 1816.

MOREL-VINDE, CHARLES-GILBERT TERRAY de, Sur la Théorie de la population, ou observations sur le système professé par M. Malthus et ses disciples, Paris, Huzard, 1829, 33 p. (2ème éd.).

NORTON, BERNARD, "La Situation intellectuelle du moment des débuts de Fisher en génétique des populations", Revue de synthèse, n°103-104, juillet-décembre 1981, p. 231-246.

NOTTER, FRIEDRICH, De qualitatibus parentum in sobolem transeuntibus, praesertim ratione rei equariae, Tubingae, Osiander, 1827.

NOTTER, FRIEDRICH, Ueber die Eigenschaften welche sich bei Menschen und Thieren von den Eltern auf die Nachkommen vererben, mit besonderer Rücksicht auf die Pferdezucht, Tübingen, Osiander, 1828 (publié au nom de Johann Daniel Hofacker).

POISSON, DENIS, « Sur la probabilité des résultats moyens des observations », Additions pour la Connaissance des tems de l'an 1827, Paris, Bachelier, 1824, p. 273-302 ; « Suite du Mémoire sur la probabilité du résultat moyen des observations », Additions pour la Connaissance des tems de l'an 1832, Paris, Bachelier, 1829, p. 3-22.

POISSON, DENIS, « Mémoire sur la proportion des naissances des filles et des garçons… », Mémoires de l'Académie royale des sciences de l'Institut de France, Paris, Firmin-Didot, IX, 1830, p. 239-308.

POISSON, DENIS ET NICOLLET, JOSEPH, « Observations relatives au nombre de naissances des deux sexes », Annuaire pour l'an 1825… par le Bureau des longitudes, Paris, Firmin-Didot, 1824, p. 98-99 (1825, p. 102-103 ; 1826, p. 92-93 ; 1832, 2ème éd., p. 98-99).

PORTER, THEODORE M., The Rise of statistical Thinking, 1820-1900, Princeton, Princeton University Press, 1986.

QUETELET, ADOLPHE, Sur l'Homme et le développement de ses facultés ou Essai de Physique sociale, Paris, Bachelier, 1835 ; Physique sociale ou Essai sur le développement des facultés de l'Homme, Bruxelles, Mucquardt - Paris, Baillière - Saint Pétersbourg, Issakoff, 1869 (Bruxelles, Académie royale de Belgique, 1997).

SADLER, MICHAEL THOMAS, The Law of Population. A Treatise, in six books, in disproof of the superfecundity of human beings, and developing the real principle of their Increase, Londres, Murray, 1830, 2 vol. (fac-simile : Londres, Routledge-Thoemmes Press, 1994).

SCHUSTER, HANS, Johann Gottfried Hoffmann als Nationalökonom. Ein Beitrag zur Geschichte der Nationalökonomie des 19. Jahrhunderts, Berlin, Walther, 1908.

SCHWEBER, LIBBY, The Assertion of disciplinary claims in demography and vital statistics France and England, 1830-1885, Ph.D. Princeton University, 1995.

SCHWEBER, LIBBY, Disciplining Statistics. Demography and Vital Statistics in France and England, 1830-1885, Durham et Londres, Duke University Press, 2006.

SPALLANZANI, LAZZARO, Expériences pour servir à l'histoire de la génération des animaux et des plantes, Genève, Chirol, 1785.

STIGLER, STEPHEN M., The History of Statistics. The Measurement of Uncertainty before 1900, Cambridge, Harvard University Press, 1986.

STIGLER, STEPHEN M., Statistics on the Table. The History of Statistical concepts and methods, Cambridge, Harvard University Press, 1999.

SÜSSMILCH, JOHANN Peter, Die göttliche Ordnung in den Veränderungen des menschlischen Geschlechts, aus der Geburt, Tod, und Fortpflanzung desselben, 1ère éd., Berlin, 1741 ; 2e éd. Berlin 1761-1762 ; trad. de l'éd. 1741 : L'Ordre divin dans les changements de l'espèce humaine démontré par la naissance, la mort et la propagation de celle-ci, Paris, Ined, 1998.

TORT, PATRICK (dir.), Dictionnaire du darwinisme et de l'évolution, Paris, P.U.F., 1996, 3 vols.VILLERME, LOUIS-RENE, « Compte-rendu de 'De l'effet de la légitimité sur le rapport des naissances de différents sexes', de Pierre Prévost (Bibliothèque universelle de Genève, octobre 1829) », Annales d'hygiène publique et de médecine légale, Paris, tome 7, 1832a, p. 453.

VILLERME, LOUIS-RENE, « Compte-rendu de 'Lettre de M. le Professeur Charles Babbage à l'Honorable M. T.-P. Courtenay, sur le rapport des deux sexes dans les naissances' (The Edinburgh Journal of Science conducted by David Brewster, 1829) », Annales d'hygiène publique et de médecine légale, Paris, tome 7, 1832b, p. 445-453.

VILLERME, LOUIS-RENE, « Compte-rendu de 'Recherches sur le rapport des deux sexes dans les naissances de M. Bickes' (Zeitung für das Gesamte Medizinalwesen, 7 février 1831) », Annales d'hygiène publique et de médecine légale, Paris, tome 7, 1832c, p. 453-459.

Capítulo 11

El modelo del dado y su influencia sobre el desarrollo de la teoría de la probabilidad. Aciertos y fracasos

MARY SOL DE MORA CHARLES
UPV/EHU, San Sebastián

Introducción

El conocido manuscrito de Cardano, que circuló por Francia antes de ser publicado póstumamente, nos ofrece la privilegiada situación de observar una investigación en su desarrollo. No es un texto acabado y corrige los errores a medida que avanza y sin retroceder nunca. Los aspectos fundamentales de su enfoque de la probabilidad han sido ya estudiados y comentados[i], como su concepto de circuito completo de posibilidades y de la mitad del circuito, que proporcionaría la probabilidad ½ y por lo tanto las condiciones para un juego justo o equitativo, donde los jugadores podrían apostar la misma cantidad sin desventaja. Estudia los juegos de dados y de cartas y en los primeros aparecen algunas variantes curiosas de la época que queremos analizar aquí.

En el texto de Cardano se ratifica una vez más la preocupación de los jugadores de todos los tiempos por impedir las trampas, en particular al lanzar los dados. Desde muy antiguo se renuncia a tirar los dados con la mano y se comienzan a utilizar cubiletes, unos pequeños vasos redondos, ordinariamente de cuerno. Con ellos se solía jugar a dados y tabas y adquieren diversos nombres y formas. Los griegos hablaban de ψηφοβολω lanzador de piedrecillas para jugar o πυργοσ (torre)[ii], que en el mundo latino se transformaría en turricula o torrecilla, orca o tonel, fimus objeto de barro, pyxis cornea o cajita (bossoli) de cuerno y sobre todo fritillus (ver figura 1). Ordinariamente no tenía fondo, era más ancho en la parte de abajo que en la superior y dentro tenía una serie de escalones inclinados que obligaban a los dados a caer de uno en otro antes de llegar a la mesa, de forma que su caída pudiera considerarse aleatoria y no provocada por un impulso voluntario del jugador. Existen numerosas citas de la antigüedad clásica en la que se menciona el fritillo y su funcionamiento: Séneca, Marcial, Juvenal, Horacio, Ausonio y otros.

Fig. 1

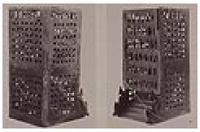
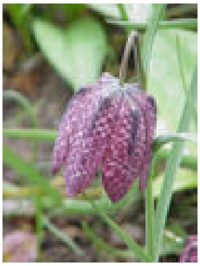
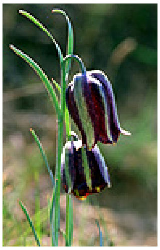


Fig. 2

Pero el destino del fritillus sería mucho más complicado, debido a algunas coincidencias azarosas. Como el propio Cardano nos asegura, los dados se solían tirar, no sobre una mesa, sino sobre un tablero que garantizaba mejor la estabilidad y nivelación del terreno de juego y en tiempos de Cardano ya se había llegado a la conclusión de que ese tablero era lo que se llamaba fritillus y no sólo él sino también un juego de características especiales que luego veremos y que se jugaba con dicho tablero, se llamaba también fritillus. Por supuesto no era en absoluto necesario que el tablero fuera de ajedrez o damas, es decir que tuviera un dibujo de cuadros o ajedrezado, pero debía ser frecuente el uso de tales tableros, porque la historia natural se inmiscuyó en el asunto y podemos encontrar flores llamadas fritillarias y también determinadas mariposas, que para mayor confusión se llaman pyrgus y cuyo aspecto parece ajedrezado o a cuadros, aunque no en todos los ejemplares. En cuanto a las flores, al menos se parecen bastante en la forma a nuestro fritillus de la figura 1:





Pyrgus bellier

Pyrgus malvae

En el capítulo VII, Cardano se explica:

De los fritillos con inclinación y los dados adulterados

> *« Coloca el tablero (tabulas) exactamente en el centro, si oscila hacia la parte adversa, favorece al oponente y por lo tanto está contra ti; igualmente si se inclina hacia tí y está desplomado en tu favor ; si el fritillo completo permanece inmóvil, la cosa está igualada. Lo mismo sucede si el fritillo recibe luz de la parte opuesta, eso es perjudicial, pues perturba la mente; por el contrario es mejor que esté contra algo oscuro. También se dice que es beneficioso si se toma posición frente a la Luna en máximo ascenso. En cuanto al dado, existen dos clases de peligro, pues todo dado, aunque sea un dado permitido, tiene un punto favorecido, ya sea por su forma o por otra causa o por casualidad, y por ello si se cambia un número grande por uno pequeño o viceversa, se comprende que la diferencia será grande. Hay otra manera, cuando el dado está adulterado porque se ha hecho su forma más redonda o más estrecha, lo cual es fácilmente apreciable, o cuando se ha dilatado en una dirección estrechándolo en los vértices contrapuestos. Por tanto, se debe hacer una prueba triple, puesto que hay tres combinaciones de vértices opuestos y eso hace que la superficie sobresalga. De modo que estos aspectos deben considerarse cuidadosamente. »*

De modo que parece identificar fritillo con tablero. Y en el capítulo XIII lo precisa aún más:

De los números compuestos, tanto hasta seis como superiores, tanto en dos dados como en tres.

> *«En dos dados, el doce y el once constan respectivamente de dos 6 y de 6 y 5. El diez de dos 5 y de 6 y 4, pero este último se puede variar de dos maneras, por lo tanto en total será la duodécima parte del circuito y la sexta de la igualdad. En el caso del nueve están el 5 y 4 y el 6 y 3, de forma que serán la novena parte del circuito y las dos novenas partes de la igualdad. El ocho se forma a partir de dos 4, de 3 y 5, y de 2 y 6. Estas cinco posibilidades forman aproximadamente la séptima parte del circuito y las dos séptimas partes de la igualdad. El siete se forma con 6 y 1, 2 y 5, 4 y 3, el total de los puntos es por lo tanto seis, la tercera parte de la igualdad y la sexta del circuito. El seis es como el ocho, el cinco como el nueve, el cuatro como el diez, el tres como el once y el dos como el doce. »*

Hasta aquí no hay problema, todo es correcto. El texto sigue así:

> *« Pero en el juego del Fritillo hay que añadir once puntos, porque se puede obtener el valor con un solo dado, así el dos se obtiene de doce maneras (con la suma 1+1 o bien con 2,1; 2,2 ; 2,3…2,6 ; 1,2 ; 3,2 ;…6,2) lo que supone dos tercios de la igualdad y un tercio del circuito. El tres por lo tanto se obtiene de trece modos, el cuatro de catorce, el cinco de quince, lo que supone diez doceavos de la igualdad y cinco doceavo del circuito completo y el seis se obtiene de dieciséis modos, lo que está muy próximo de la igualdad. »*

Es decir 16/36 o lo que es lo mismo, (8/18)(2/2) = (8/9)(1/2). A partir del siete, ya no se puede obtener el valor con un sólo dado y Cardano no lo escribe, o bien la edición lo ha omitido. Cardano lo resume en esta tabla, que también parece haber sido mal transcrita por los editores:

Consensus sortis in duabus Aleis.

2	12	1 uno	3	11	2 dos	4	10	3 tres	Aequal.
5	9	4 cuatro	6	8	5 cinco	7 seis	**8**	18 dieciocho en el Fritillo (igualdad)	Ad Frit.

En las Suertes se juega con la suma de los puntos y no vale considerar cada dado por separado o en conjunto, a voluntad, como en cambio sí sucede en el Fritillo (considerado aquí como un juego diferente). De ese modo, en la tabla, el ocho debería tener cinco posibilidades en las Suertes y lo mismo en el Fritillo. Se ha dejado llevar por la progresión de las posibilidades del dos al seis, añadiendo siempre 11 puntos, sin darse cuenta de que a partir del siete ya no se puede hacer así.

Estas serían las posibilidades para el Fritillo: El dos se puede hacer también con un sólo dado, es decir, de once maneras más 1 debido a dos dados,

(1,2),(2,1),(2,2),(2,3),(3,2),(2,4),(4,2),(2,5),(5,2),(2,6),(6,2), (1,1).

(12/36)(2/2) = (2/3) (1/2)

El 3 se puede hacer de la misma manera, es decir, 11+2:

(13/36)(2/2)= (13/18) (1/2) = (6,5/9) (1/2).

El cuatro: 11+3

(14/36)(2/2) = (7/9) (1/2).

El cinco: 11+4

(15/36)(2/2) = (5/6)(1/2) = (7,5/9)(1/2).

El seis: 11+5

(16/36)(2/2) = (8/9) (1/2) muy cerca de la igualdad.

A continuación aparece la tabla para tres dados, tanto en las Suertes (ahora llamadas Aleis) como en el Fritillo:

Consensus sortis in tribus Aleis tum Frit.

Sortis			Fritilli	
			3	115
3	18	1 uno	4	125
4	17	3 tres	5	126
5	16	6 seis	6	133
6	15	10 diez	7	33
7	14	15 quince	8	36
8	13	21 veintiuno	9	37

9	12	25 venticinco	10	36
10	11	27 veintisiete	11	38
			12	26

Del mismo modo que en las Suertes, aquí hay números como:

13 que es 21 (veintiuno)
14 que es 15(quince)

Para las suertes, todos los valores son correctos:

El 3 y el 18:

(1, 1, 1) y (6, 6, 6)

El cuatro y el diecisiete:

(1,1,2), (2,1,1), (1,2,1) y (5,6,6), (6,5,6), (6,6,5) Etc.

Pero en el caso del fritillus, hay muchos problemas:

El valor para el tres es correcto:

(1,2, x), (3, x, y)

(1, 1, 1) y sus permutaciones:

(1,1,2) ~~(1,1,2)~~ (1,2,1) (2,1,2) (1,2,2) ~~(1,2,2)~~ (3,1,2) (1,3,2) (1,2,3)	7
(4,1,2) (1,4,2) (1,2,4) (5,1,2) (1,5,2) (1,2,5) (6,1,2) (1,6,2) (1,2,6)	9
~~(1,2,1)~~ (2,1,1) (~~2,1,1~~) (2,2,1) (~~2,2,1~~) (~~2,1,2~~) (3,2,1) (2,3,1) (2,1,3)	5
(4,2,1) (2,4,1) (2,1,4) (5,2,1) (2,5,1) (2,1,5) (6,2,1) (2,6,1) (2,1,6)	9
(1,3,1) (3,1,1) (~~3,1,1~~) (~~2,3,1~~) (~~3,2,1~~) (~~3,1,2~~) (3,3,1) (~~3,3,1~~) (3,1,3)	4
(4,3,1) (3,4,1)(3,1,4) (5,3,1) (3,5,1) (3,1,5) (6,3,1) (3,6,1) (3,1,6)	9
(1,1,3) (~~1,1,3~~) (~~1,3,1~~) (~~2,1,3~~) (~~1,2,3~~) (~~1,3,2~~) (~~3,1,3~~) (1,3,3) (~~1,3,3~~)	2
(4,1,3) (1,4,3) (1,3,4) (5,1,3) (1,5,3) (1,3,5) (6,1,3) (1,6,3) (1,3,6)	9
(~~1,3,2~~) (~~3,1,2~~) (~~3,2,1~~) (2,3,2) (3,2,2) (~~3,2,2~~) (3,3,2) (~~3,3,2~~)(3,2,3)	4
(4,3,2)(3,4,2)(3,2,4) (5,3,2) (3,5,2)(3,2,5) (6,3,2)(3,6,2)(3,2,6)	9
(~~1,2,3~~) (~~2,1,3~~) (~~2,3,1~~) (2,2,3)(~~2,2,3~~)(~~2,3,2~~) (~~3,2,3~~)(2,3,3)(~~2,3,3~~)	2
(4,2,3)(2,4,3)(2,3,4) (5,2,3)(2,5,3)(2,3,5) (6,2,3)(2,6,3)(2,3,6)	9
(~~1,3,4~~)(~~3,1,4~~)(~~3,4,1~~) (~~2,3,4~~)(~~3,2,4~~)(~~3,4,2~~) (3,3,4)(~~3,3,4~~)(3,4,3)	2
(4,3,4)(3,4,4)(~~3,4,4~~) (5,3,4)(3,5,4)(3,4,5) (6,3,4)(3,6,4)(3,4,6)	8
(~~1,4,3~~)(~~4,1,3~~)(~~4,3,1~~) (~~2,4,3~~)(~~4,2,3~~)(~~4,3,2~~) (~~3,4,3~~)(4,3,3)(~~4,3,3~~)	1
(4,4,3)(~~4,4,3~~)(~~4,3,4~~) (5,4,3)(4,5,3)(4,3,5) (6,4,3)(4,6,3)(4,3,6)	7
(~~1,3,5~~)(~~3,1,5~~)(~~3,5,1~~) (~~2,3,5~~)(~~3,2,5~~)(~~3,5,2~~) (3,3,5)(~~3,3,5~~)(3,5,3)	2
(~~4,3,5~~)(~~3,4,5~~)(~~3,5,4~~) (5,3,5)(3,5,5)(~~3,5,5~~) (6,3,5)(3,6,5)(3,5,6)	5
(~~1,5,3~~)(~~5,1,3~~)(~~5,3,1~~) (~~2,5,3~~)(~~5,2,3~~)(~~5,3,2~~) (~~3,5,3~~)(5,3,3)(~~5,3,3~~)	1
(~~4,5,3~~)(~~5,4,3~~)(~~5,3,4~~) (5,5,3)(~~5,5,3~~)(~~5,3,5~~) (6,5,3)(5,6,3)(5,3,6)	4
(~~1,3,6~~)(~~3,1,6~~)(~~3,6,1~~) (~~2,3,6~~)(~~3,2,6~~)(~~3,6,2~~) (3,3,6)(~~3,3,6~~)(3,6,3)	2
(~~4,3,6~~)(~~3,4,6~~)(~~3,6,4~~) (~~5,3,6~~)(~~3,5,6~~)(~~3,6,5~~) (6,3,6)(3,6,6)(~~3,6,6~~)	2
(~~1,6,3~~)(~~6,1,3~~)(~~6,3,1~~) (~~2,6,3~~)(~~6,2,3~~)(~~6,3,2~~) (~~3,6,3~~)(6,3,3)(~~6,3,3~~)	1
(~~4,6,3~~)(~~6,4,3~~)(~~6,3,4~~) (~~5,6,3~~)(~~6,5,3~~)(~~6,3,5~~) (6,6,3)(~~6,6,3~~)(~~6,3,6~~)	1

Más (1, 1, 1), suman 115 posibilidades, tal como Cardano calculó, al parecer escribiéndolas todas como hemos hecho aquí.

Para el cuatro, tendríamos

A: (1, 1, 2), (2, 1, 1), (1,2,1) y además, con dos dados :

B :(1, 3, x) y sus permutaciones con repetición: 30

C :(2, 2, x) y sus permutaciones con repetición: 18

D: (4, x, y) y lo mismo con un solo dado. Menos los casos repetidos. Se aplica el teorema:

p(A) +p(B) + p(C)+p(D) – p(A y B) –p(A y C) –p(A y D)

– p(B y C) – p(B y D) – p(C y D) + p(A y B y C)

+p(A y B y D) + p(B y C y D) - p(A y B y C y D)

El resultado es pues 3 + 30 + 18 + 82 = 133, el resultado que él da para el seis.

Para el cinco sería: (1, 4, x), (2, 3, x), (5, x, y), (1,2,2)

(1, 1, 3) y todas sus permutaciones.

En total serían 6 + 60 + 85 = 151.

Y para el seis:

(2, 4, x) Aquí cuenta seis y son 30;

(3, 3, x) Aquí cuenta tres y son 18

(3, 3, x) Aquí cuenta tres y son 18

(1, 5, x) Aquí cuenta seis y son 30

(6, x, y), (2,2,2), (1,2,3), (1,1,4)

En total serían: 10 + 78 + 79 = 167.

Para el siete: ya sabemos que no se puede hacer con un solo dado. Con dos dados sí:

(1, 6, x) Cuenta seis, son 30; (2, 5, x) lo mismo; (3, 4, x) lo mismo,

(2, 2 ,3), (1, 3, 3), (1,1,5), (1,2,4)

En resumen serían: 15+ 90 = 105. De todas formas aquí se obtiene menos que la igualdad, como Cardano observó.

Para el ocho:

(4, 4, x) Cardano cuenta tres, es correcto.

(3, 5, x) Cuenta seis, pero son 30, si hemos de seguir la pauta establecida para el cuatro.

(2, 6, x) Lo mismo.

(1,3,4), (1,2,5), (1,1,6), (2,2,4), (2,3,3).

En resumen serían: 21 + 63 = 84 casos y no treinta y seis

Para el nueve:

(3, 6, x) Cuenta seis pero son 30
(4, 5, x) Lo mismo.
(1,4,4), (1,3,5), (1,2,6), (2,2,5), (2,3,4), (3,3,3).

En resumen serían: 25 + 60 = 85 y no treinta y siete.

Para el diez:

(5, 5, x) son 3 posibilidades
(4, 6, x) son 30, y no 6, como arriba
(1,3,6), (1,4,5), (2,2,6), (2,3,5), (2,4,4), (3,3,4).

En resumen serían: 27 + 33 = 60 y no treinta y seis

Para el once:

(6, 5, x) De nuevo cuenta seis pero son 30
(1,4,6), (2,3,6), (2,4,5), (3,3,5), (3,4,4), (5,1,5)

En resumen serían: 27 + 30 = 57 posibilidades y no treinta y ocho.

Y para el doce:

(6, 6, x) Cuenta uno pero son 18 casos.
(1,5,6), (2,4,6), (2,5,5), (3,3,6), (3,4,5), (4,4,4)

En resumen serían: 25+18 = 43 y no veintiséis.

Después del doce, ya no se puede hacer con dos dados, así que los valores son los de las suertes hasta 18. Esto es correcto.

A continuación afirma: « *Además un punto tiene 108, dos puntos tienen 11* ». Este error se subsanará más adelante, en el capítulo XIV, donde se considera la probabilidad de obtener un punto determinado al tirar tres dados (o un dado tres veces), con el primero de ellos, o con el segundo, o con el tercero o bien con dos o con los tres dados, que produce el valor 91/216. Este es su razonamiento, que nos proporciona un valioso ejemplo de la progresión de su pensamiento hacia los resultados correctos:

De los puntos combinados

« En el caso de dos dados, debemos entrar en un razonamiento del tipo siguiente, que el punto 1 tiene once tiradas (favorables) y el punto 2 igualmente y el 3, y así todos los puntos singulares, pero el punto 1 y el 2 no tienen veintidós casos sino veinte. Pues el 1 tiene once y el 2 nueve. Y así, si se añade el 3, no serán veintinueve ni treinta y uno, sino veintisiete, y los números de las tiradas que se obtienen de este modo se pueden ver en la tabla:

20 (11+9)	*11*
27 (20+7)	*9*
32 (27+5)	*7*
35 (32+3)	*5*
36 (35+1)	*3*

Así, si se consideran todas las tiradas se obtienen treinta y seis, pues con ello el circuito se hace perfecto y es necesario que todas las tiradas contengan algún punto, de modo que se completa el número del circuito. Sin embargo si alguien dice, quiero un 1 o un 3, tú sabes que son veintisiete tiradas favorables, y como en el circuito son treinta y seis, las tiradas restantes en las que estos puntos no salen serán nueve, y la proporción por tanto será de 3 a 1. En cuatro tiradas con la misma fortuna, los puntos 1, 2 o 3 saldrán tres veces, y sólo habrá una tirada en la que no esté en ninguno de ellos; sin embargo, si apuesta tres ases el que espera los puntos 1 a 3, y el otro apuesta uno, el primero ganará tres veces y ganará tres ases, el otro una vez y ganará tres ases, por tanto el circuito de cuatro tiradas será siempre equitativo. Así pues este es el razonamiento para jugar en condiciones iguales, pues si otro apuesta más, jugará en condiciones injustas y con pérdidas, y si apuesta menos, con ventaja ».

Analiza a continuación los casos de 1, 2, 3 o 4, y de dos 1 o dos 2, y pasa al caso de tres dados:

«Los mismos razonamientos se observan en tres dados, tanto en los puntos simples como en los compuestos, y proponemos como anteriormente que las tiradas para un punto son 108, por lo tanto será necesario establecer seis términos de los cuales el máximo será 108 y los restantes guardarán distancia igual entre sí respecto a aquel y tales que completen 216, como se ve en la tabla:

91 (para un punto)	30
61 (para dos puntos)	24
37 (para tres puntos)	18
19 (para cuatro puntos)	12
7 (para cinco puntos)	6.
1 (para seis puntos)	
216 (total de casos posibles)	

Pero ningún punto obtiene la mitad de todo el circuito, pues la proporción es de 91 a 125, o si la invertimos, muy próxima de 25 a 18, mayor por lo tanto de 4 a 3. Luego el que así apostase a que no salía (el punto) ganará, en tanto que en siete tiradas no haya salido, y si apuesta 4, ganará todavía 3. Del mismo modo se considera en los restantes casos, pues es evidente que con dos dados los incrementos son iguales. Pero para tres tenemos un exceso igual, como se muestra en la tabla. »

Vemos pues cómo el fritillo es en realidad el teorema de la unión de sucesos. Por otra parte, cuando intenta la generalización a varias tiradas no queda tan claro:

« Pero si fueran necesarias dos tiradas, los multiplicaremos entre sí, y si fueran tres o cuatro, haríamos lo mismo y en proporción a los números así obtenidos tendríamos que hacer la comparación. Así si fuera necesario que alguien sacara un 1 dos veces, en ese caso sabes que el número correspondiente es 91, y el resto es 125.

Así, multiplicamos cada uno de esos números por sí mismo y obtenemos 8281 y 15625, y la proporción es aproximadamente de 2 a 1. Si apostase el doble, contendería bajo condiciones injustas, aunque en opinión de algunos, la condición de que alguien ofrezca doble apuesta sería mejor. Por lo tanto, en tres tiradas sucesivas, si se necesita (en todas ellas al menos) un 1, la proporción sería de 753.571 a 1,953.125, proporción que es próxima a 5 a 2, aunque algo mayor. »

De hecho es 2,59. No obstante, esa operación no es correcta, porque implica que si $a + b = c$, también $a^n + b^n = c^n$ lo cual como sabemos no es cierto. Si queremos que un suceso determinado se repita, y en su repetición es independiente de los resultados anteriores o posteriores, la fórmula adecuada es la de intersección de probabilidades que aplica de facto Cardano: $p(A \text{ y } B \text{ y } C) = p(A) . p(B). p(C) = (91/216). (91/216). (91/216) = 753571/10077696 = 1/13{,}37$. Y el resto no es $(125)^3$ sino $10077696 – 753571 = 9324125$ es decir, 1 a 123,7. Lo cual muestra lo fácil que es distraerse al sacar conclusiones.

Finalmente, en el capítulo XXX, Cardano aclara aún más la interpretación que del fritillum se hacía en su época:

De los juegos de azar entre los antiguos

« …Así, contra todas esos trucos han imaginado a la orca (tonel), por su semejanza con el pez, pues se ve que devora el dado como la orca devora otros peces más pequeños. Persius /Saty.3):

No ser engañado por el cuello de la angosta orca

Pomponio, el poeta de Bolonia:

Mientras contemplaba la orca he perdido el pequeño dado.

Horacio llama a esta pyxis (cajita) un pyrgus (torre), utilizando una palabra griega, cuando dice (Saty.9, Sermon. 2:

Pon las tabas en el pyrgus.

Pues ponían en ella no sólo los astrágalos, sino los dados; es de uso constante en Bolonia, pero no en Milán. Marcial la llama turricula (torrecilla) y dice en el Apophoreta sobre la turrícula:

> La mano impía intenta atratar y lanzar los dados, si los lanza a través de mí, siempre obtiene lo que desea.

Este juego de dados se modifica para ser jugado con el fritillus (pues ese no es el pyrgo sino un tablero de juego), digámoslo así, no quiero discutir por las palabras.»

---

[i] Véase M.S. de Mora, Los inicios de la teoría de la probabilidad, UPV/EHU, 1989, p.22-57.
[ii] Hofmann, Johann Jacob (1635-1706): Lexicon Universale, Historiam Sacram Et Profanam Omnis aevi, Leiden, 1698.voz Pyrgus.

Capítulo 12

El libro de los dados de Alfonso X
Su relación con el cálculo de probabilidades

JESÚS BASULTO SANTOS
JOSÉ ANTONIO CAMÚÑEZ RUIZ
Universidad de Sevilla
CÉSAR BORDONS ALBA
I. E. S. Punta de Verde

Introducción

El rey Alfonso X es merecidamente conocido con el sobrenombre de Sabio. Fueron muchas las áreas de conocimiento que de una forma u otra recabaron la atención y el interés del rey. Podemos comenzar recordando su notable aportación literaria, poética y musical; sus trabajos sobre legislación, historiografía y ciencias naturales –en un sentido muy amplio, que incluiría campos que hoy consideraríamos pseudo científicos y mágicos-; y podemos terminar mencionando su interés por los juegos.

Primero el rey Sabio se había interesado principalmente por actividades recreativas que podríamos calificar de atléticas y que tenían una vinculación directa con la necesidad de estar bien preparados para la guerra, función social primordial de reyes, príncipes y nobles en la sociedad medieval. Estos ejercicios se podían hacer en solitario, en compañía e incluso a nivel competitivo. Se trata de caminar, montar a caballo, arrojar lanzas, disparar flechas con un arco apuntando a un blanco a pie o a caballo y cazar. Sabemos que Fernando III disfrutaba oyendo música y que incluso se atrevía a cantar, además era un reconocido jugador de juegos de mesa. Su hijo heredó y desarrolló estas aficiones –además de la de la lectura-, por lo cual no debe sorprendernos que diera cabida en su amplio programa cultural a un libro sobre los juegos.

El Libro del ajedrez, dados y tablas se terminó de escribir en Sevilla en el año 1283, un año antes de la muerte de Alfonso X. Se trata de un manuscrito de 98 folios de pergamino encuadernado en piel. Las hojas son de gran tamaño, de 40 x 28 cm. En la actualidad el original se encuentra en la Biblioteca del Monasterio de El Escorial, a donde fue llevado en 1591 por Felipe II desde la Capilla Real de Granada. Esta obra es un ejemplo perfecto de la labor intelectual del Rey Sabio, que consistió fundamentalmente en ser el gran transmisor de la cultura y la sabiduría oriental a la Península Ibérica y desde aquí a todo Occidente.

El Libro del ajedrez, dados y tablas consta de siete partes, separadas por folios en blanco. No debe sorprendernos la utilización de este número tan querido por nuestro rey. Ya hemos hecho referencia al interés real por los asuntos mágicos, y en este campo el número siete tiene una interesante simbología a la que el rey Alfonso aludirá en más de una ocasión. Las siete partes son las siguientes:

1. El libro del ajedrez. (folios 1r a 64r)
2. El libro de los dados. (folios 65r a 71v)
3. El libro de las tablas. (folios 72r a 80r)
4. Ajedrez y tablas decimales. (El grant açedrez) (folios 81r a 85v)
5. Otros juegos. (Açedrez de los quatro tiempos) (folios 87r a 89v)
6. El libro del alquerque. (folios 91r a 93v)
7. Juegos astronómicos. (folios 95r a 97v)

Casi al comienzo del texto se recoge una antigua leyenda situada en la India en torno a la polémica sobre la causalidad de los acontecimientos, que se ilustra comparando las ventajas respectivas de los juegos del ajedrez, los dados y las tablas. El sabio partidario de la uentura argumenta que nada vale el seso, que lo decisivo es la suerte, que llega al hombre para proporcionarle pro o danno, y respalda su razonamiento con el ejemplo del juego de los dados.

> *"Segunt cuenta en las historias antiguas en India la mayor ouo un Rey que amaua mucho los sabios e tenielos siempre consigo e fazieles mucho amenudo razonar sobre los fechos que nascien de las cosas. E destos auie y tres que tenien sennas razones. El uno dizie que mas ualie seso que uentura. Ca el que uiuie por el seso; fazie sus cosas ordenadamientre e aun que perdiesse; que no auie y culpa. Pues que fazie lo quel conuinie. Ell otro dizie que mas ualie uentura que seso, ca si uentura ouiesse de perder o de ganar; que por ningun seso que ouiesse; non podrie estorcer dello. El tercero dizie que era meior qui pudiesse ueuir tomando delo uno e delo al, ca esto era cordura, ca en el seso quanto meior era; tanto auie y mayor cuydado como se pudiesse fazer complidamientre. E otrossi en la uentura quanto mayor era; que tanto auie y mayor peligro por que no es cosa cierta. Mas la cordura derecha era; tomar del seso aquello que entendiesse omne que mas su pro fuesse, e dela uentura guardarse omne de su danno lo mas que pudiesse e ayudarse della en lo que fuesse su pro.*

Libro del ajedrez, dados y tablas, folio 2v

E desque ouieron dichas sus razones much affincadas; mandoles el Rey quel aduxiesse ende cadauno muestra de prueua daquello que dizien, e dioles plazo qual le demandaron, e ellos fueron se e cataron sus libros, cadauno segunt su razon. E quando llego el plazo, uinieron cada unos antel Rey con su muestra. E el que tenie razon del seso, troxo el acedrex con sus iuegos mostrando que el que mayor seso ouiesse, e estudiesse aperçebudo podrie uencer all otro. E el segundo que tenie la razon dela uentura troxo los dados mostrando que no ualie nada el seso si no la uentura, segunt parescie por la suerte llegando el omne por ella a pro o a danno. El tercero que dizie que era meior tomar delo uno e delo al, troxo el tablero con sus tablas contadas e puestas en sus casas ordenadamientre, e con sus dados, que las mouiessen pora iugar ..." (folios 1v y 2r)

Y concluye, ya en el folio siguiente:

"E por que el acedrex es mas assessegado iuego e onrrando que los dados nin las tablas, fabla en este libro primeramientre del."

Posteriormente el rey se sentirá obligado a justificar por qué trata de los dados en segundo lugar. No por que los considere en segunda posición en cuanto a interés y merecimiento, sino por una cuestión práctica, ya que los dados son necesarios para jugar a las tablas, entretenimiento que goza de mayor aprobación real. La obra de Alfonso X está siempre planteada con un claro afán didáctico, por lo tanto lo racional es hablar primero de los dados, explicar cómo se fabrican y cómo se juega con ellos, antes de pasar a otros juegos que requieren de su utilización.

"Pues que de los iuegos del açedrex que se iuegan por seso auemos ya fablado lo más complidamientre que pudiemos queremos agora aqui contar de los iuegos de los dados por dos razones. La una por que la contienda de los sabios, segund mostramos en el comienço del libro, fue entre seso e uentura qual era meior. E desto dio cada uno so muestra al rey.
El primero del seso, por los iuegos del açedrex. E el segundo de la auentura: por los dados.
La otra por que maguer las tablas son mayor cosa e mas apersonada que los dados por que ellas non se pueden iogar a menos dellos, conuiene que fablemos dellos primeramientre." (Folio 65r)

Recordemos aquí lo que en el Libro de las tablas se dice sobre los dados.

"Los dados a mester por fuerça que ayan las tablas. Ca bien assi como el cuerpo non se podrie mouer sin los pies, assi ellas non se mouerien sin ellos pora fazaer ningun ieugo, ca por fuerça derecha segunt los puntos de los dados an ellas de iogar en aquellas casas que son señaladas pora ellas…" (folio 72v)

Otra prueba más de la menor consideración en que se tenía al juego de los dados la encontramos en que en las miniaturas que ilustran el Libro de ajedrez y el Libro de las tablas aparecen con frecuencia mujeres y niños practicándolos, así como personajes reales, ya que socialmente no estaban bien considerados. En el Libro de los dados aparecen mujeres en las miniaturas de los folios 68r y 68v, donde están sirviendo bebidas a los jugadores, y jugando en los folios 69r y 70r.

El Libro de los dados consta de siete folios, del folio 65r al 71v. Catorce páginas en total, doce de ellas enriquecidas con miniaturas alusivas al juego de los dados. En la primera de ellas tenemos una miniatura donde aparece el rey dictando el texto a un copista que se halla situado a su derecha, mientras que a su izquierda se acumulan cinco personajes, cuatro de los cuales parecen dirigirse al rey para solicitar su intervención en alguno de los muchos problemas que ocasionaba el juego de los dados. Esta primera miniatura (65r), al igual que las que dan comienzo a los libros de ajedrez (1r) y de tablas (72r) tiene un enorme interés para comprender el sistema de trabajo de la corte de Alfonso X y el propio universo mental del rey. Ana Domínguez Rodríguez[1] señaló acertadamente que en estas miniaturas, como en otras similares de otras obras alfonsíes que incluyen ilustraciones donde se representa al rey enfrascado en su labor intelectual, podemos encontrar superpuestos dos elementos: el primero imperial-mayestático, resaltado por los atributos reales (corona, trono, lujosa vestimenta) y el solemne escenario; y el segundo hermético-filosófico, señalado por la postura del rey con el dedo índice de la mano derecha enhiesto (postura filosófica) señalando un libro.

Libro del ajedrez, dados y tablas, folio 65r

Ya desde el lejano año de 1915 Antonio García Solalinde[2] nos desveló cuál era el sistema de trabajo del equipo del rey Sabio, cuál era la intervención de Alfonso X en la redacción de sus obras. La participación del rey era fundamental: el rey ordenaba, impulsaba y dirigía el trabajo y cuando ya estaba acabado lo corregía. Recordemos lo que el propio rey escribía en la General Estoria:

> *"El rey faze un libro, non porquel escriva con sus manos, mas porque compone la razones del, e las enmienda e yegua e endereça, e muestra la manera de cómo se deven fazer, e desi escríbelas qi el manda, pero dezimos por esta razon que el faze el libro."*

El sistema de trabajo del libro de los dados seguiría este mismo esquema. El rey se ocuparía de proporcionar todo el material necesario, haciendo que trajeran al equipo de redactores toda la información necesaria en forma de manuscritos o libros sobre el tema. El rey era quien seleccionaba a los autores que iban a formar el equipo de trabajo y asignaba las diversas funciones de cada uno. Elaboraba un boceto de lo que quería que hicieran sus colaboradores y supervisaba todo el trabajo, introduciendo continuamente correcciones. Finalmente el rey tenía que aprobar la versión definitiva.

---

[1] A. DOMÍNGUEZ, "El libro de los juegos y la miniatura alfonsí", 45. En la edición facsímil que del Libro del ajedrez, dados y tablas coeditaron en 1987 ediciones Poniente de Madrid y Vincent García Editores, S.A. de Valencia.

[2] A. GARCÍA SOLALINDE, (1915). Intervención de Alfonso X en la redacción de sus obras, Revista de Filología Española, II, 283-288.

En la segunda miniatura del Libro de los dados (65v) podemos observar el proceso de fabricación de los dados. Hay que volver a llamar la atención sobre la estructura didáctica de la obra alfonsí: primero el rey dictando el libro, en segundo lugar la fabricación de los dados y solo después ya se podrá empezar a hablar de los diversos juegos de dados. Ya en esta segunda miniatura el rey nos advierte de la naturaleza peligrosa del juego de los dados. En la derecha de la ilustración se nos representa a un jugador que acude al taller a comprar unos dados, ha perdido la camisa y los zapatos y además en su cara se refleja su desgracia. En el texto, a la vez que se va explicando cómo han de elaborarse los dados se alerta sobre la existencia de dados "engañosos":

> *"E dezimos que han de seer tres figuras quadradas de seys cantos eguales tamanno el uno como el otro en grandez e en egualdad de la quadra, ca ssi en otra manera fuesse no caerie tan bien de una parte como de otra, e serie enganno mas que uentura. E por ende esta es la una de las maneras de enganno como diremos adelante: con que fazen los dados engannosos aquellos que quieren engannar con ellos.*
> *E a de auer en estas seys quadras en cada una dellas: puntos puestos en esta guisa.*
> *En la una seys e en la otra cinco, en la otra quatro, en la otra tres, en la otra dos, e en la otra uno, assi que uengan en cada un dado ueynti un punto, de manera que uengan en los tres dados sessenta e tres puntos. E deuen seer puestos los puntos en esta guisa, so la faz del seys: el as, e so el cinco: el dos, e so el quatro: el tria.*
> *E estos dados pueden seer fechos de fuste, o de piedra, o de huesso, o de todo metal; mas sennaladamientre, son meiores de huesso el mas pesado que fallaren, que de otra cosa ninguna, e mas ygualmientre e mas llanos caen doquier que los echen."* (folios 65r y 65v)

Libro del ajedrez, dados y tablas, folio 66v

El resto de las miniaturas son de gran valor sociológico ya que en ellas se representan diversas escenas relacionadas con el juego de los dados. Ya en la primera de éstas (66r) observamos la presencia de un puñal sobre el tablero de juego, en la siguiente (67r) encontramos a dos jugadores enzarzados en una pelea y al resto de los participantes descamisados. La violencia irá aumentando hasta que nos encontremos en la miniatura que ilustra el folio 70v a dos jugadores dándose puñaladas. Son frecuentes las escenas de apuestas y de empeño de bienes (67v y 69r).

También se describen algunas escenas tranquilas relacionadas con el juego de los dados, como la del folio 68r, donde un músico aparece amenizando con su laúd a los jugadores y una mujer ofrece una copa a uno de los jugadores; en el Libro del ajedrez este gesto tenía un valor simbólico, se ofrecía la copa al jugador que ganaba. Muy interesante resulta la ilustración del folio 71v, que es el que pone fin al Libro de los dados, en ella observamos a dos grupos de cuatro jugadores, pero lo más destacable es el marco en donde se desarrolla la partida, en el acogedor jardín –con palmeras y naranjos- de una casa acomodada. Desgraciadamente las escenas apacibles son la excepción, parece que este tipo de actividad generaba una gran conflictividad, como podemos corroborar con la lectura del Ordenamiento de la tafurerias, que había mandado redactar el propio Alfonso X unos años antes, en 1276.

Libro del ajedrez, dados y tablas, folio 71v

Resulta altamente interesante constatar que la palabra tafur –actualmente tahúr- comienza a utilizarse en castellano precisamente en estos momentos, en la segunda mitad del siglo XIII. Señala Corominas[3] que se trata de una voz común a todas las lenguas romances de Francia y de la Península Ibérica y que parece haber designado primero a los componentes de una tropa auxiliar de las Cruzadas, que se dedicaban al saqueo y al merodeo. Se trataba de una muchedumbre andrajosa y hambrienta que aunque se dedicaba fundamentalmente al merodeo, también atacaba con temible valor, pero que vivía de forma miserable, extendiéndose incluso el rumor de que habían devorado cadáveres sarracenos. El origen del vocablo parece estar en el armenio "taphur", que significa abandonado, desnudo, vagabundo. La primera utilización del término en España se constata en un documento de 1260 fechado en Sevilla, donde se utiliza como nombre propio, se habla de un tal Pedro Royz Taffur. El término se utilizará repetidamente en la Gran conquista de Ultramar y también lo empleará más adelante Juan Ruiz, con la acepción que ya se había generalizado de jugador vicioso y fullero.

No es el momento de resaltar la gran trascendencia que tuvo la obra legislativa de Alfonso X, baste recordar solamente que es la primera vez que se utiliza la lengua vernácula para estos asuntos; que engloba derecho civil y canónico; que obedece al deseo del rey Sabio de fortalecer el poder real frente a una poderosa y peligrosa nobleza; y que por último –y no menos importante-, proporciona información fundamental a los estudiosos de la sociedad del momento, en nuestro caso a los interesados en el juego.

Centrándonos en la legislación sobre el juego hemos de tener presente que los jugadores estaban especialmente mal considerados por la sociedad en general y por las autoridades, especialmente las eclesiásticas, debido a que frecuentemente los juegos de dados terminaban

[3] J. COROMINAS. Diccionario crítico etimológico castellano e hispánico. Editorial Gredos. Madrid, 1980. Tomo V, página 377

con blasfemias que vulneraban el tercero de los mandamientos de la iglesia: no tomar el nombre de Dios en vano[4]. Sabemos, tanto por los códigos legales como por la literatura didáctica de la época que cuando un jugador perdía su partida -y su apuesta- era habitual que desahogase su frustración con insultos a Dios, a la Santa Virgen María y a todos los santos, incluso se han recogido testimonios de jugadores desairados que escupían sobre la cruz.

En el Espéculo, que data de mediados de la década de 1250 se contempla que si alguna persona que forma parte del séquito real regenta casas de juego o truca los dados para engañar a sus compañeros de juego será desterrado. También se establece que este tipo de personas no puede comparecer como testigo ante un tribunal, ya que su palabra no tiene valor.

Las Siete Partidas (1256-1265) van a dedicar más atención legislativa a los jugadores y a las apuestas, parece como si el juego y los problemas que éste acarrea estuviesen en expansión en la Castilla de esos años. Continúa la incapacitación de los jugadores como testigos, ya que se les equipara a ladrones, bandidos y proxenetas. Se resalta la prohibición del juego a los clérigos, incluso se les prohíbe estar en compañía de jugadores. Se contemplan los delitos de los estafadores que utilizan dados trucados, o de los que aprovechan del jaleo que se crea -o que ellos mismos fingen-, para desplumar a los incautos que se concentran en torno a la partida de dados. Todo un título se dedica al asunto de la blasfemia, que ya hemos comentado que esta muy vinculado a este mundo.

La legislación del Espéculo y de las Siete Partidas con respecto al juego obedece fundamentalmente a consideraciones de tipo moral y a preocupaciones de seguridad pública, de evitar conflictos. Pero a partir de la revuelta mudéjar (1264-1266) y del reinicio de la guerra con los musulmanes el rey va a modificar su actitud ante el problema del juego, adoptando una política completamente prohibicionista que se justifica por la preocupación real ante la ruina financiera y moral que el juego provocaba en muchas personas. Esta proscripción del juego va reforzada con el establecimiento de durísimas sanciones económicas para los que lo practicasen.

Como ocurrirá en muchos otros casos similares, pronto se puso de manifiesto que resultaba imposible hacer cumplir los buenos deseos reales, el juego era algo muy arraigado en la sociedad y Alfonso X, paulatinamente, va a ir adoptando una política más realista. Pocos años después va a conceder, primero a la ciudad de Sevilla y luego a la de Murcia una carta de privilegios en las que les permite el establecimiento de una casa de juego, regulada por los oficiales del rey y sometida a los correspondientes impuestos. También se permite a los hombres buenos de la ciudad jugar en su propia casa y en cualquier otro lugar, sin ser molestados por ello por la autoridad. Esta última concesión tiene interés de tipo sociológico, para la mentalidad del rey no son lo mismo desde el punto de vista moral un hombre bueno que un hombre del común, estos últimos serían mucho más vulnerables al vicio y necesitarían del proteccionismo paternal real.

[4] R. A. MACDONALD. Libro de las tahurerías: a special code of law, concerning. gambling, drawn up by Maestro Roldán at the command of Alfonso X of Castile. Madison Hispanic Seminary of Medieval Studies, 1995

Libro del ajedrez, dados y tablas, folio 72v

A mediados de la década de los setenta el rey Sabio ha llegado ya al convencimiento de que el problema del juego debe ser afrontado de una manera más comprensiva, además, si se establece un sistema de concesiones de lugares de juego se podrían conseguir unos ingresos para el fisco real, que en esos momentos se encontraba en una situación especialmente delicada. Con estas premisas, Alfonso X encargó a Maestro Roldán la elaboración de un código legal que regulara el juego. Piedra central del nuevo sistema van a ser las tafurerías, lugares que recibían una concesión real para poder practicar en ellos el juego legalmente. Estas tafurerías se van a localizar preferentemente en el ámbito urbano. El encargado de regentar estos nuevos lugares de juegos es denominado en el Ordenamiento de las Tahurerías "tablajero".

Los 12 juegos del libro de los dados

El Libro de los dados contiene los siguientes doce juegos: de Mayores, de Menores, de Tanto en Uno como en Dos, Triga, juego de Azar, juego de Marlota, juego de Riffa, juego de Par con As, Panquist, juego de Medio Azar, juego de Azar Pujado y juego de Guirguiesca.

El juego de Triga es el que va a generar los juegos de Azar, Marlota, Panquist y Guirguiesca. En todos se utilizan tres dados excepto el juego de Guirguiesca que usa dos. Todos estos juegos dividen las puntuaciones totales de los tres dados en dos conjuntos llamados azar y suerte, que son disjuntos y complementarios. Además se usan los conceptos de soçobra[5], encuentro y, en el caso de que las suertes repartidas entre los dos jugadores sean distintas, uno de los jugadores lanzará sucesivamente los datos hasta que aparezca la suerte de uno de ellos que le hará ganar.

El juego de Azar es una complicación del juego de Triga, en cambio en juego de Marlota es una simplificación del Triga. El juego de Panquist es una complicación del juego de

[5] Esta palabra soçobra la traducimos por zozobra que significa en nuestro caso "cara del dado opuesta a la que se considera". María Moliner, (1977). Diccionario del Uso del Español. Editorial Gredos. Madrid.

Marlota. Por último, el juego de Guirguiesca añade, a las apuestas que lo jugadores realizan al comienzo del juego, la posibilidad de hacer un envite en el proceso del juego.

Otros tipos de juegos son el de Tanto en uno como dos, el juego de Riffa y el de Par con As. Estos juegos hacen uso de turnos en los lanzamientos y buscan que gane el jugador que antes lance algún tipo de coincidencia entre los tres dados o que logre una mayor puntación total al lanzar los tres dados bajo ciertas condiciones.

Por último, los dos juegos de Medio Azar y de Azar Pujado son juegos donde los jugadores pueden ganar dinero cuando, al lanzar los dados, ocurren puntuaciones del conjunto Azar en varias partes del juego. En estos juegos no se busca la probabilidad de que gane el juego uno de los jugadores, como ocurre en los otros juegos anteriores. Lo que se busca es aumentar la cantidad de dinero que se puede ganar. De ahí que evaluaremos estos juegos a partir de calcular las cantidad esperada de dinero que puede ganar uno de los jugadores.

Muchos de los juegos que recoge el manuscrito de Alfonso X aparecerán modificados, pero manteniendo ciertas estructuras, en las obras de los matemáticos que vivieron en los siglos XVI, XVII y XVIII, como Cardano, Galileo, Huygens, Leibniz, Montmort, De Moivre, J. Bernoulli, entre otros.

Otras manuscritos que recogen cálculos con dados, juegos de dados o juegos de dados con tablas son: el Vetula, fechado entre 1266-1269, los Carmina Burana, siglo XIII, que contiene miniaturas como el manuscrito de Alfonso X, los dos grandes trabajos latinos de Bonus Socius (antes de 1300) y Civis Bononiae (1400-1450)[6] que contienen problemas sobre tablas y problemas con dados relacionados con el cálculo de las probabilidades[7], por último está el manuscrito denominado "Royal 13 A XVIII-1300" que contiene un texto latino llamado "Ludi ad tabulas".

A partir de aquí vamos comenzar a describir cada uno de los juegos recogidos por el manuscrito de Alfonso X.

El juego de mayores y el de tanto en uno como en dos

El manuscrito recoge el texto siguiente,

El iuego de mayores e de tanto en uno como en dos

> *"El primer iuego de los que usan los omnes, el que mas puntos echare, que gane; e este iuego que llaman a mayores.*
> *Todas las otras maneras de iuegos que a en ellos son posturas que pusieron los omnes entre ssi que son juegos departidos.*
> *Assi como qui echasse menos puntos que ganasse.*
> *O tanto en ell uno como en los dos, que es en esta manera que si dixiere en el un dado seys, que diga en los otros dos cinco e as o quatro e dos e ternas. E si dixiere en el uno cinco, que diga en los otros dos quatro e as, o tres e dos. E si dixiere en ell uno quatro, que diga en los otros tria e as, o dos dos. E si dixiere en ell otro tres, que diga en los otros*

[6] Murray, H.J.R. (1913). History of Ches. Oxford. Capítulo VII.
[7] Murray, H.J.R. (1978). A History of Board-Games. Other than Chess. Hacker art Books Inc. New York. Capítulo 6.

dos e as. E si dixiere en ell otro dos, que diga en los otros amas as." *(Folio 65v)*

Libro del ajedrez, dados y tablas, folio 66r

En esta miniatura aparecen dos grupos de jugadores jugando con tres dados a los juegos De Mayores y de Tanto en Uno como en Dos. Vemos que la suma de las puntuaciones de estos juegos es 21=15+6, es decir, un dado con seis puntuaciones del tablero de la izquierda es la soçobra de uno de los dados con una puntuación del tablero de la derecha; igual ocurre con el dado de tres puntuaciones del tablero de la izquierda que es la soçobra del dado de cuatro puntuaciones del tablero de la derecha. Globalmente diremos, más adelante, que la puntuación de los tres dados de la miniatura de la izquierda es la soçobra de la puntuación de los tres dados de la derecha. También se ve un puñal en el tablero de la izquierda.

Veamos a continuación nuestra interpretación de los juegos.

El Libro de los dados recoge, en este primer punto, tres juegos. El primero, que llama De mayores [1], es un juego de dados donde gana el jugador que, al lanzar los dados, saca la puntuación más alta. En este primer juego es fácil de ver que la probabilidad de que gane uno de los jugadores es igual a 0,5. No importa quién comience lanzando los dados. El segundo, que llama De menores [2], es un juego de dados donde gana el jugador que, al lanzar los dados, saca la puntuación más pequeña.

A continuación define los juegos de partido como los que permiten que un jugador haga envites o apuestas de acuerdo con los otros jugadores. En un juego de partido el jugador no busca la mayor puntuación, como en el juego De mayores, ya que también puede ganar el juego si lanza otras puntuaciones. Un ejemplo es el siguiente juego llamado Tanto en uno como en dos [3]

Este tercer juego, que el Libro llama Tanto en uno como en dos, se juega con tres dados. El juego consiste en que uno de los jugadores, que vamos a llamar G1, lanza los tres dados y si ocurre una de las siguientes combinaciones con repetición: 6-5-1, 6-4-2, 6-3-3, 5-4-1, 5-3-2, 4-3-1, 4-2-2, 3-2-1 ó 1-1-1, gana. Si no logra lanzar una de las anteriores combinaciones, entonces deberá lanzar los tres dados el otro jugador que llamaremos G2. Si G2 tampoco lanza una de las anteriores combinaciones, entonces lanzará los dados el jugador G1 y así sucesivamente hasta que alguno de los jugadores lance una de las combinaciones anteriores, lo que le hará ganar el juego.

Vamos a denominar configuraciones a las combinaciones con repetición. Recogemos en la Tabla 3-1 las probabilidades de cada una de las configuraciones de este juego de Tanto en uno como en dos.

Tabla 3-1

Configuraciones	Probabilidades de cada Configuración
6-5-1, 6-4-2, 5-4-1, 5-3-2, 4-3-1, 3-2-1	6/216
6-3-3, 4-2-2, 2-1-1	3/216

Vemos que el juego finaliza cuando ocurre que los puntos de un dado coinciden con la suma de los puntos de los otros dos dados. También se observa que cuando una configuración tiene puntos distintos, como 6-5-1, entonces ocurre según 6 posibilidades. Por ejemplo, la configuración 6-5-1 puede ocurrir según las siguientes permutaciones con repetición: (6,5,1), (6,1,5), (5, 6,1), (5,1,6), (1,6,5) y (1,5,6).

En este juego de Tanto en uno como en dos, no se sortea qué jugador debe comenzar tirando los tres dados. Veremos que en otros juegos del Libro de los Dados se sortea quién debe comenzar lanzando los dados (en el manuscrito se dice "hacer batalla). La probabilidad de que gane este juego el jugador G1, en el caso de que lance el primero, es igual a 0,5581. Por lo tanto, el jugador que lanza primero tiene una ventaja de 0,0581 (5,81%), un valor pequeño que justifica el que no se sortee quién debe lanzar primero los tres dados. Vemos también que los jugadores se turnan en el lanzamiento de los dados.

Juegos que hacen uso de turnos podemos encontrarlos en el libro de Huygens (1629-1696) "De Ratiociniis in Ludo Aleae" de 1657, el primer libro de cálculo de probabilidades. Una traducción al castellano de la edición inglesa puede verse en la Tesis Doctoral del profesor Camúñez[8]. En este libro traducido por "El Cálculo en los Juegos de Azar" Huygens resuelve un juego de dados con turnos en su Proposición XIV. Recordemos este juego de Huygens.

Si otro jugador y yo lanzamos por turno 2 dados con la condición de que yo habré ganado lanzando 7 puntos y él cuando haya lanzado 6, y si le dejo lanzar el primero, encontrar la relación de mi chance a la suya. La solución de este juego la deduce Huygens, siendo la relación de las chances de 31 a 30.

Vemos que a pesar que el otro jugador comienza tirando los dados, su probabilidad de ganar es menor, 0,4918, que la mía, 0,5082. Esto último se debe a que, aunque comienza tirando los dos dados, no es suficiente para disminuir la ventaja que yo tengo con mi puntuación 7, con seis casos favorables, frente a la puntuación suya de 6, con cinco casos favorables.

El juego de Triga [4]

El manuscrito recoge el texto siguiente,

Este es el iuego de la triga

> *"Otro iuego ay que llaman triga que se iuega en esta manera, que si omne iuega con otro, e lança primeramientre par en los tres dados, o quinze puntos o dizeseys o dizesiete o dizeocho, o la soçobra destos que son seys, cinco e quatro e tres que gana; e estas suertes todas son llamadas trigas e pueden uenir en esta manera. Los dizeocho puntos, senas alterz. Los dizesiete, senas cinco. Los dizeseys senas quatro, e quinas seis. Los quinze senas tria, e seys cinco, e quatro e quinas alterz."*

[8] "La Probabilidad y la Estadística en el Período 1654-1670. Sus Antecedentes en el Renacimiento Italiano". Memoria de Tesis Doctoral. 2004. Universidad de Sevilla.

Otra manera de triga

“Otrossi los seys puntos pueden uenir en esta manera quatro amas as o tres dos e as, o dos dos alterz. Los cinco puntos tria e amas as o dos dos as. Los quatro dos e amas as. Los tres amas as alterz. Otrossi par en todos los dados pueden uenir en esta manera, senas alterz, quinas alterz, quadernas alterz, dos alterz, amas as alterz.”

Otra manera de triga

“En otra manera puede omne lançar, en que no avra ninguna destas suertes, que fata aqui diximos e sera triga, como si omne tomare pora ssi, siete puntos o ocho o nueve o diez o onze o doze o treze o catorze. E ell otro con qui ell iogare lançare aquella mis/ (66r)ma suerte, esta sera triga e ganara el qui primero tomo suerte. Et ssi por auentura non lançare la suerte del otro e tomare otra pora ssi, conuerna que lance tantas vezes fasta que qual quiere dellos acierte en alguna destas suertes. E lançando la su suerte sera triga e ganara. Et si lançare la del otro, otrossi será triga e perdera.
E estas suertes pueden venir en tantas maneras. Los siete puntos cinco amas as o quatro dos e as o tres e dos dos o ternas as. Los ocho puntos seys amas as, o cinco dos e as, o quatro dos dos, o ternas dos. Los nueve puntos seys dos e as o cinco tria e as, o cinco dos dos, o quadernas as, o quatro tres e dos o ternas alterz. Los diez puntos seys tria e as, o seys dos dos, o cinco quatro e as, o cinco tres e dos, o quadernas dos o ternas quatro. Los onze puntos seys quatro e as, o seys tres e dos, o cinco quatro e dos, o quinas as, o ternas cinco, o quadernas tria. Los doze puntos seys cinco e as, seys quatro e dos, o seys e ternas, o cinco quatro e tria o quinas dos o quadernas alterz. Los treze puntos senas as, o seys quatro e tria, o seys cinco e dos, o quinas tria o quadernas cinco. Los catorze puntos senas dos, o seys cinco e tria, o seys e quadernas, o quinas quatro.
E en tantas maneras como desuso auemos dicho, pueden venir suertes, en los dados, e no en mas.” (Folio 66v)

Libro del ajedrez, dados y tablas, folio 67r

La miniatura asociada a este juego de Triga muestra las consecuencias que el juego tiene sobre la moral de los jugadores. Vemos que uno de los jugadores ha lanzado una puntuación 16 que es un azar, que como sabemos es muy poco probable (seis en 216).

Veamos a continuación nuestra interpretación de este juego de Triga.

Primera parte del juego de Triga [4]

La primera parte del juego de triga comienza en: Este es el iuego de la triga. Es un juego donde dos jugadores, que llamaremos G1 y G2, lanzan tres dados. Este juego considera, primero, las puntuaciones, suma de los puntos observados en los tres dados, igual a {15, 16, 17, 18}. Para las quince puntuaciones produce las configuraciones: 6-6-3, 6-5-4 y 5-5-5; para dieciséis puntuaciones: 6-6-4 y 6-5-5; para diecisiete: 6-6-5 y, por último, para dieciocho: 6-6-6. A continuación añade a estas puntuaciones sus soçobras, que son {6, 5, 4, 3}, respectivamente. Al conjunto de las puntuaciones {15, 16, 17, 18, 6, 5, 4, 3} vamos a llamarlo conjunto Azar.

La palabra soçobra significa el complemento de la suma de las puntuaciones observadas al lanzar los tres dados respecto de 21 (donde, 1+2+3+4+5+6=21). Por ejemplo, la soçobra de 15 es 6 porque 15+6=21. También, la soçobra de una configuración (combinación con repetición), por ejemplo, 6-6-4, es 1-1-3, es decir, tomamos el complemento respecto de 7 de cada valor numérico de la configuración (por ejemplo, del valor 6, su complemento es 7-6=1).

Ahora, esta Primera Parte de Triga consiste en que uno de los jugadores, por ejemplo, G1, comienza lanzando los tres dados, y si lanza una puntuación del conjunto Azar gana el juego. Si G1 no lanza una puntuación del conjunto Azar, entonces lanzará una puntuación del complementario del conjunto azar (que veremos en la Segunda Parte del juego) y, a continuación, pasará el turno al jugador, G2, que ganará el juego si lanza una puntuación del conjunto Azar. Si G2 no lanza una puntuación del conjunto Azar, entonces lanzará una puntuación del complementario del conjunto azar (que veremos en la Segunda Parte del juego)

Las configuraciones de los puntos {6, 5, 4, 3} que son las soçobras, respectivamente, de las puntuaciones {15, 16, 17, 18}, son descritas en la parte: Otra manera de triga. Estas configuraciones son: 4-1-1, 3-2-1, 2-2-2 para las seis puntuaciones; 3-1-1, 2-2-1 para las cinco puntuaciones; 2-1-1 para las cuatro puntuaciones y 1-1-1 para las tres puntuaciones.

En el manuscrito se habla de lanzar par en los tres dados, es decir, que al lanzar los dados salgan los mismos puntos en los tres. Notemos que el conjunto Azar recoge las configuraciones: 1-1-1, 2-2-2, 5-5-5 y 6-6-6. En el manuscrito, al hablar de par, incluye la configuración 4-4-4 pero no recoge la configuración 3-3-3. Nuestra opinión es que debemos no considerar, en esta primera parte del Juego de Triga, la configuración 4-4-4 si queremos ser coherente con la otra parte del juego de triga que vemos a continuación.

Segunda parte del juego de Triga [4]

La segunda parte del juego de triga comienza en la frase repetida: Otra manera de triga.

En esta segunda parte, las configuraciones de interés son ahora las que generan las puntuaciones del conjunto {14, 13, 12, 11, 10, 9, 8, 7}, que vamos a llamar conjunto Suerte. Las configuraciones correspondientes a este conjunto Suerte son: 6-6-2, 6-5-3, 6-4-4, 5-5-4 para catorce puntuaciones; 6-6-1, 6-5-2, 6-4-3, 5-5-3, 4-4-5 para trece puntuaciones; 6-5-1, 6-4-2, 6-3-3, 5-5-2, 5-4-3, 4-4-4 para doce puntuaciones; 6-4-1, 6-3-2, 5-5-1, 5-4-2, 5-3-3, 4-4-3 para once puntuaciones; 6-3-1, 6-2-2, 5-4-1, 5-3-2, 4-4-2, 4-3-3 para diez puntuaciones; 6-2-1, 5-3-1, 5-2-2, 4-4-1, 4-3-2, 3-3-3 para nueve puntuaciones; 6-1-1, 5-2-1, 4-3-1 (la configuración 4-3-1 no es recogida en el manuscrito), 4-2-2, 3-3-2 para ocho puntuaciones; 5-1-1, 4-2-1, 3-3-1, 2-2-3 para siete puntuaciones.

Ahora, si el jugador G1, en su primer lanzamiento, saca una puntuación suerte, ésta será su suerte. A continuación debe lanzar los dados el jugador G2. Si el jugador G2 lanza una puntuación suerte, ésta será su suerte. Si las suertes son iguales, entonces gana el juego el jugador que primero tomó su suerte (que es el que comenzó lanzando los dados). En otro caso, si las suertes lanzadas por los jugadores son diferentes, entonces se lanzarán los tres dados hasta que venga la suerte del jugador, el jugador G1, que le hace ganar, o la suerte del jugador, G2, que sería entonces el ganador.

En esta Segunda Parte del Juego de Triga se afirma al comienzo que "En otra manera puede omne lançar, en que no avra ninguna destas suertes, que fata aqui dixiemos…", esta frase refuerza lo que anteriormente dijimos sobre quitar la configuración 4-4-4 del Primer Juego de Triga.

Recogemos en la Tabla 4-1 los casos favorables de cada una de las configuraciones de este juego de Triga.

Tabla 4-1

Puntos	Configuraciones						Casos Favorables
18	6-6-6						1
17	6-6-5						3
16	6-6-4	6-5-5					6
15	6-6-3	6-5-4	5-5-5				10
14	6-6-2	6-5-3	6-4-4	5-5-4			15
13	6-6-1	6-5-2	6-4-3	5-5-3	4-4-5		21
12	6-5-1	6-4-2	6-3-3	5-5-2	5-4-3	4-4-4	25
11	6-4-1	6-3-2	5-5-1	5-4-2	5-3-3	4-4-3	27
10	6-3-1	6-2-2	5-4-1	5-3-2	4-4-2	4-3-3	27
9	6-2-1	5-3-1	5-2-2	4-4-1	4-3-2	3-3-3	25
8	6-1-1	5-2-1	4-3-1	4-2-2	3-3-2		21
7	5-1-1	4-2-1	3-3-1	2-2-3			15
6	4-1-1	3-2-1	2-2-2				10
5	3-1-1	2-2-1					6
4	2-1-1						3
3	1-1-1						1

De la Tabla 4-1 podemos ver que las puntuaciones del conjunto Azar (el color azul o sombreado), de la primera parte del Juego de Triga, son menos probables que las puntuaciones del conjunto Suerte (de la Segunda Parte del Juego), lo que explicaría por qué al conjunto {15, 16, 17, 18, 6, 5, 4, 3} se le llama conjunto azar. También vemos que una puntuación y su soçobra tienen la misma probabilidad, con lo que la palabra soçobra recogería el resultado empírico de que las puntuaciones que suman 21 deben aparecer, al lanzar muchas veces los dados, con las mismas frecuencias.

Las configuraciones que hemos recogido en la Tabla 4-1, de los juegos de Triga, son las mismas que las recogidas en el poema de Vetula fechado a mediados del siglo XIII. Una traducción al castellano de la parte de este poema de Vetula puede verse en la Tesis Doctoral de J. A. Camúñez (2004). La notación que usa el poema de Vetula para las Configuraciones es, por ejemplo, 411, mientras que nosotros hemos utilizado la notación 4-1-1. La traducción de esta parte del poema de Vetula, en latín e inglés, puede verse en Bellhouse (2000).

Este método de representar las combinaciones con repetición, las configuraciones, será retomado por Cardano (1501-1576). Cardano, en su obra Liber de Ludo Aleae (1564) utiliza el mismo método que el del Libro de los Dados. En el Capítulo Décimo-Tercero (ver la traducción de Marisol De Mora Charles (1989)) escribe: "En dos dados, el doce y el once constan respectivamente de dos 6 y de 6 y 5. El diez de dos 5 y de 6 y 4, pero este último se puede variar de dos maneras, etc. ". (In duabus Aleis duodecim, & vndecim constan eadem ratione qua bis, sex, atque sex, & quinque. Decem autem ex bis quinque, & sex, & quatuor, hoc autem variatur dupliciter, etc.).Vemos que si en la combinación con repetición 6 y 4 considera sus dos permutaciones con repetición, en la combinación 6 y 5, las omite. Esto último, ocurre con las puntuaciones nueve, ocho, siete, etc.

En el caso de tres dados, Cardano calcula, para cada puntuación, el total de sus permutaciones con repetición. Para las puntuaciones cinco y seis, Cardano nos aporta, además, las correspondientes permutaciones con repetición. En cambio, en el poema de Vetuta se calculan las permutaciones con repetición y los totales para cada una de las puntuaciones. Estos totales los recogemos en la última columna de la Tabla 4-2.

Tabla 4-2

3	18	configuraciones de puntos sobre los dados	1	formas de caer (Cadentia)	1
4	17	configuraciones de puntos sobre los dados	1	formas de caer (Cadentia)	3
5	16	configuraciones de puntos sobre los dados	2	formas de caer (Cadentia)	6
6	15	configuraciones de puntos sobre los dados	3	formas de caer (Cadentia)	10
7	14	configuraciones de puntos sobre los dados	4	formas de caer (Cadentia)	15
8	13	configuraciones de puntos sobre los dados	5	formas de caer (Cadentia)	21
9	12	configuraciones de puntos sobre los dados	6	formas de caer (Cadentia)	25
10	11	configuraciones de puntos sobre los dados	6	formas de caer (Cadentia)	27

En esta Tabla 4-2 del poema se recoge en las dos primeras columnas cada puntuación con su soçobra, como, por ejemplo, las puntaciones 3 y 18. En la última columna se recogen las chances o el total de las permutaciones con repetición para cada una de las puntuaciones, así, por ejemplo, la puntuación 4 ocurre 3 veces frente al total de 216 ya que la puntación 4 genera las siguientes permutaciones con repetición o chances: (2,1,1), (1,2,1) y (1,1,2). También, en la columna cuatro de la Tabla 4-2, el poema ha recogido el total de configuraciones (combinaciones con repetición) correspondientes a cada una de las puntuaciones, por ejemplo, para la puntuación 4 sólo hay una configuración. Incluimos la palabra latina "cadentia" que se refiere a todas las formas que tienen los dados de salir.

Cardano, en el capítulo XIII de su obra, elabora una tabla que denomina "Consesus sortis in duabus Aleis" donde recoge las combinaciones con repetición para las puntuaciones que resultan al lanzar dos dados y sus correspondientes totales de las permutaciones con repetición. Una siguiente tabla con el título "Suertes y Fritillos para juegos que usan tres dados" (traducción de Marisol de Mora Charles), incluye las columnas primera, segunda y última de la Tabla 4-2.

D. R. Bellhouse (2004) ha propuesto que el libro de Cardano pueda provenir del poema de Vetula, aunque ha señalado algunas diferencias entre ellos. Una de las diferencias que Bellhouse comenta es entre el uso de "sortis" por Cardano, frente al uso de "cadentia", permutaciones con repetición, en el poema de Vetula. Que Cardano use "sortis", al igual que el uso de "suerte" en el Libro de Alfonso X, nos lleva a inclinarnos a que el Libro de Cardano pueda tener también relación con el Libro de los Dados.

Veamos a continuación que la probabilidad de que gane el juego el jugador G1, que comienza lanzando los dados, es igual a 0,5604. Para llegar a esta probabilidad debemos recordar que el jugador G1 puede ganar en los siguientes lanzamientos: (1) Gana si lanza una puntuación azar en el primer lanzamiento; (2) Gana si lanza una puntuación suerte en su primer lanzamiento y el jugador G2 lanza, en su primer lanzamiento (segundo lanzamiento del juego), la misma surte; (3) Si las suertes repartidas entre los dos jugadores son distintas, gana el jugador G1 si al lanzar los tres dados sucesivamente sale antes la suerte de G1.

Primero, la probabilidad de que G1 gane el juego cuando saca en su primer lanzamiento una puntuación azar es 0,185. Segundo, debemos calcular las probabilidades de que G1 lance una puntuación suerte y a continuación G2 lance otra puntuación suerte, que puede ser la misma. Estas probabilidades son el producto de las probabilidades de cada una de las suertes, por ejemplo, si la suerte de G1 es 8 y la de G2 7, entonces la probabilidad de que ocurran ambas suertes es,

$$P_r\left[8,7\right]=\frac{CF\left(8\right)*CF\left(7\right)}{216*216}=\frac{21*15}{46656}=0,06751. \text{ [1]}$$

Con esta fórmula podemos calcular la probabilidad de que gane este juego de Triga el jugador G1 cuando las suertes son iguales, sumando las probabilidades para las suertes idénticas. Aplicando la fórmula [1] obtenemos una probabilidad de 0,0866.

Por último, el jugador G1 puede ganar el juego cuando las suertes repartidas son distintas. Ahora calculamos la probabilidad, condicionada a cada par de suertes distintas de cada jugador, de que al lanzar los dados sucesivamente llegue primero la suerte del jugador G1 antes que la del jugador G2. Una fórmula de estás probabilidades para cada par de surtes distintas es la siguiente,

$$P_r\left[S\left(G_1\right)\middle|S\left(G_1\right),S\left(G_2\right)\right]=\frac{CF\left(S\left(G_1\right)\right)}{CF\left(S\left(G_1\right)\right)+CF\left(G_2\right)}, \text{ [2]}$$

donde S(G_1) y S(G_2) son las suerte de G1 y G2 respectivamente y CF(S(G_1)) y CF((S(G_2)) son los casos favorables de las suertes de G1 y G2, respectivamente. Por ejemplo, para las suerte 8 y 7, de G1 y G2, respectivamente, la fórmula [2] vale 21/(21+15)=0,58333.

Con las fórmulas [1] y [2] podemos calcular la probabilidad de que gane el juego de Triga cuando las suertes dos distintas, así un cálculo laborioso nos lleva a una probabilidad igual a 0,28867.

Si sumamos las tres probabilidades, obtenemos el valor de 0,5604. Vemos que el jugador que comienza lanzando los dados tiene una cierta ventaja, frente al otro jugador, de ganar este juego de Triga. Esta ventaja viene de que el jugador G1 gana el juego cuando saca, en el primer lanzamiento, una puntuación azar o si el jugador G2 lanza la misma suerte del jugador G1. Si modificamos este juego de Triga, haciendo que gane el jugador G2 cuando las suertes repartidas son iguales, un cálculo directo nos lleva a que la probabilidad de que G1 gane este juego de Triga es ahora 0,4738, una probabilidad que favorece al jugador G2. Otra alternativa sería introducir turnos cuando las suertes repartidas son distintas y además que comience lanzando los dados el jugador G2, en este caso la probabilidad de ganar este juego de Triga el jugador G1 es 0,53569.

Si hacemos los cálculos anteriores bajo el supuesto de que la combinaciones con repetición, las configuraciones, son equiprobables, obtenemos que el jugador G1 tiene ahora una probabilidad de 0,5672 de ganar, que está próxima a la verdadera, 0,5604. Suponer que las configuraciones son equiprobables puede verse en varios manuscritos sobre juegos de azar de Leibniz (Mora Charles, 1992) y en el libro de Cotton (1674). Probabilidades tan cercanas no ayudarían a distinguir empíricamente entre la hipótesis verdadera, de las permutaciones con repetición, frente a la falsa, de las combinaciones con repetición. Sólo un análisis combinatorio a la manera del poema de Vetula o de Galileo justificará la hipótesis de las permutaciones con repetición.

El juego de Azar [5]

El manuscrito recoge el texto siguiente,

El iuego que llaman azar.

> *"Otra manera ay de iuego de dados que llaman azar que se iuega en esta guisa. El qui primero ouiere de lançar los dados si lançare xv. puntos o dizeseys o dizesiete o dizeocho, o las soçobras destas suertes, que son seys o cinco o quatro o tres: gana. E qual quiere destas suertes en qual quier manera que uengan segundo los otros iuegos que desuso dixiemos es llamado azar.*
> *E si por auentura no lança ninguno destos azares primeramientre, e da all otro por suerte una daquellas que son de seys puntos a arriba o de quinze ayuso, en qual quiere manera que pueda uenir, segundo en los otros iuegos dixiemos que vinien. E depues destas lançare alguna de las suertes que aqui dixiemos que son azar, esta suerte sera llamada, reazar, e perdera aquel que primero lançare. E otrossi si por auentura no lançare esta suerte, que se torna en reazar, e tomare pora si una de las otras suertes que son de seys puntos a arriba o de quinze ayuso en qual quiere manera que uenga. Conuerna que lancen tantas uegadas fasta que uenga una destas suertes o la suya por que gana, o la dell otro por que pierde, saluo ende si tomare aquella misma suerte que dio all otro: que serie llamada encuentro. E conuernie que tornassen a alançar como de cabo. E como quier que uiniesse alguna de las suertes que son llamadas azar o reazar, e entre tanto que uinie una daquellas que amos auien tomado pora ssi: non ganarie ninguno dellos por ella nin perderie fasta que se partiesse por las suertes, assi como desuso dize." (Folio 67r)*

Veamos a continuación nuestra interpretación de este juego de Azar.

Llamaremos G1 al jugador que lanza los tres dados y al otro G2. Si en el primer lanzamiento sale una puntuación azar, es decir, {3,4,5,6,18,17,16,15}, entonces G1 gana el juego (el total apostado); si sale la puntuación que llama suerte, es decir, {7,8,9,10,14,13,12,11}, entonces le corresponde al jugador G2 esta puntuación suerte (por ejemplo, si salen 7 puntos en este primer lanzamiento, entonces la suerte de G2 es de 7 puntos).

En el segundo lanzamiento puede salir una puntuación azar, lo que hace perder al jugador G1, o una puntuación suerte, por ejemplo, 14 puntos, entonces esta puntuación suerte le corresponderá al jugador G1.

Si las suertes repartidas, por ejemplo, 7 puntos de G1 y 14 puntos de G2, son distintas, como ocurre en este ejemplo, entonces el jugador G1 deberá lanzar los tres dados hasta que venga la suerte de G1, que le hace ganar, o la suerte de G2, que hace perder a G1 y ganar a G2. Si ocurre una puntuación azar, esto no hace ganar o perder a los jugadores. Por último, cuando las suertes repartidas son iguales (se dice que hay "encuentro"), entonces el juego debe comenzar de nuevo.

El este juego se llama reazar a la puntuación azar que sale en el segundo lanzamiento. La palabra rezar no se puede interpretar como "que vuelve a salir otra vez la puntuación azar" porque si el jugador G1 ha sacado una puntuación azar en segundo lanzamiento, que le hace perder el juego, necesariamente no debe haber salido una puntuación azar en lanzamiento primero.

La diferencia entre los juegos de Azar y Triga está en la solución que se da cuando ocurre encuentro, ya que si en el juego de Triga se hace ganar al jugador que comienza lanzando los dados, en este juego de Azar se propone que el juego comience de nuevo, lo que hace aumentar la duración del juego.

En la Ley XL, del Ordenamiento de la Tahurerías, se habla de este juego de Azar, diciendo que se trata de un juego de "paradas". La palabra "paradas" recoge las distintas apuestas que los jugadores realizaban al comienzo o a lo largo del juego.

En este juego de Azar, ambos jugadores han apostado una misma cantidad antes de comenzar los lanzamientos de los tres dados. En la Ley se habla de que los jugadores apostaban seis maravedíes cada uno. En cambio cuando uno de los jugadores hace un envite, como veremos en el juego de guirguesca, la Ley habla de apostar seis o cinco maravedíes (estas apuestas serían hechas en las Casas de Juegos Reales).

La probabilidad de que gane este juego de Azar el jugador G1 es 0,5187 (Basulto et al, 2006). Al estar esta probabilidad cercana a 0,5, explicaría por qué en este juego no se sortea quién debe lanzar los dados (en el manuscrito, "hacer batalla"). En nuestro trabajo estudiamos la relación que tiene este juego de Azar con los juegos de Azar de los autores clásicos de la Historia del Cálculo de Probabilidades, como Pierre Rémond de Montmort (1714), Abraham de Moivre (1718) y Jame Bernoulli (1713).

Entre 1613 y 1623 Galileo escribió, poco antes de llegar a Florencia, un trabajo con el título: Sopra le Scorpete dei Dadi (Sobre un descubrimiento concerniente a los dados) cuya traducción al castellano puede verse en la Tesis Doctoral del profesor J. A. Camúñez. Este trabajo fue publicado en 1718, cuando aparecen las obras completas de Galileo. El trabajo que realiza Galileo fue ordenado por el gran duque de la Toscaza. En palabras de Galileo: "Ahora yo, para complacer a quien me ha pedido que muestre lo que se me ocurra sobre tal problema, expondré mis ideas, con la esperanza no sólo de solucionar este problema, sino también de abrir el camino hacia una comprensión exacta de las razones por las que todos los detalles del juego han estado dispuestos y ajustados con gran cuidado y juicio". Galileo, que va directamente a calcular las probabilidades de obtener la suma de las puntuaciones que resultan al lanzar tres dados, afirma: "El hecho de que ciertos números de un juego de dados sean más ventajosos que otros tiene una razón muy obvia, que es, que algunos se consiguen más fácilmente y con más frecuencia que otros, lo cual depende de su capacidad de ser compuestos de una mayor variedad de números". Una vez calculados, en una tabla, las chances asociadas a cada una de las posibles sumas, Galileo termina su trabajo con la siguientes palabras: “Y con esta tabla, cualquiera que entienda el juego puede medir muy ajustadamente todas las ventajas, por pequeñas que puedan ser, del zara, del incontri, y de alguna otra regla especial y condición observada en este juego”. Vemos que se trataría de un juego de Azar, donde la palabra italiana zara es nuestra palabra azar y la palabra italiana “incontri” corresponde a “encuentro”.

El juego de Marlota [6]

El manuscrito recoge el texto siguiente,

Este es el iuego de marlota.

> *“Otro iuego ay de dados que llaman marlota en que no a azar nin reazar nin triga. E iuegasse por suerte partida desta guisa. El que lançare los dados a de dar suerte al otro con que iogare. E las suertes que puede dar o tomar pora ssi a este iuego son estas siete o ocho o nueve, o diez o onze o doze o treze o catorze, en qual quiere manera que uengan segundo en los otros juegos dixiemos que pueden uenir, e si lançare de catorze a arriba o de siete ayuso, no es suerte pora al uno ni pora al otro ante conuerna que lance tantas ueces fasta que de suerte destas sobredichas a aquel con que iogare, e tome otra pora ssi, e destas suertes a de seer la primera daquel con que iogare, e la otra suya. E depues que las suertes fueren partidas en esta guisa, a de lançar tanto fasta que uenga la suya o la del otro, e assi lançando la suya gana e lançando la dell otro pierde.” (Folio 67v)*

Veamos a continuación nuestra interpretación de este juego de Marlota.

Este es un juego donde el jugador que lanza los tres dados debe lanzarlos hasta que salga una puntuación del conjunto suerte {7, 8, 9, 10, 11, 12, 13, 14}, por ejemplo, 7 puntos. Esta puntuación suerte será del jugador G2. A continuación el jugador G1 seguirá lanzando los tres dados hasta que salga una suerte “distinta” (en el manuscrito se habla de que “tome otra”) de la suerte dada a G2, por ejemplo 13. Esta última suerte será la del jugador G1. Con estas suertes repartidas, el jugador lanzara sucesivamente los tres dados hasta que venga la suerte de G1, que le hace ganar, o que venga la suerte de G2, que le hace perder (ganando el jugador G2).

En este juego no se habla de “encuentro” porque se exige que las suertes repartidas sean distintas, así el juego no se repite y se evita entrar en un juego “infinito” como en el juego anterior de Azar. Este hecho de que finalice el juego cuando ocurre “encuentro” es lo que distingue a este juego de Marlota del juego de Azar.

También vemos que en este juego de Marlota se han eliminado las puntuaciones de azar y reazar, lo que evita que el juego termine pronto porque salga una puntuación azar en el primer lanzamiento, o que salga reazar en el segundo lanzamiento, como en los juegos de Triga y Azar. Por lo tanto queda explicado que el manuscrito afirme: “Otro iuego ay de dados que llaman marlota en que no a azar nin reazar nin triga”.

En la Ley XL, del Ordenamiento de la Tahurerías, se habla de este juego de Marlota, diciendo que se trata de un juego de “paradas”. También llama a este juego “marveta”. Este nuevo nombre para el juego de marlota puede ser un error de los que copiaron estas Leyes.

Veamos ahora las probabilidades asociadas a las suertes repartidas entre los jugadores G1 y G2. Las probabilidades que tiene de ganar el jugador G1, una vez se han repartido las suertes son calculadas por la fórmula [2].

Si queremos conocer la probabilidad de que el jugador G1 gane este juego antes de haber lanzado las suertes, es decir, desde el comienzo, necesitamos dar una interpretación de la primera parte de este juego.

Recogemos en la Tabla 6-1 las puntuaciones suertes y, para cada una de ellas, el total de permutaciones con repetición que genera cada puntuación suerte.

Tabla 6-1

Puntuaciones Suertes	7	8	9	10	11	12	13	14	Total
Frecuencias	15	21	25	27	27	25	21	15	176

Si ahora consideramos una urna que contiene las puntuaciones surtes, {7, 8, 9, 10, 11, 12, 13, 14}, repetidas según las frecuencias absolutas de la Tabla 6-1, entonces podemos interpretar la parte primera de este juego de Marlota de la siguiente forma: el jugador G1 debe sacar dos números sin reposición de esta supuesta urna: 1) sacará un número, por ejemplo, un 7, que será la suerte del jugador G2, y 2) quitará todos los números repetidos del número que haya sacado, por ejemplo, si ha salido un 7 quitara los 15 sietes de la urna, y a continuación sacará otro número, por ejemplo, la puntuación 8, que será su suerte. Ahora la probabilidad de que ocurran las suertes 7 y 8 es,

$$P_r[7,8]=\frac{CF(7)*CF(8)}{176*(176-CF(7))}=\frac{15*21}{176*(176-15)}=0,01111. \quad [3]$$

A partir de las fórmulas [2], que usamos en el juego de Triga, y [3] obtenemos que el jugador G1 tiene una probabilidad de 0,4948 de ganar este juego de Marlota. Al estar esta última probabilidad cercana a 0.5 explica por qué no se hace "batalla" en este juego. Si hacemos los cálculos anteriores bajo el supuesto de que las combinaciones con repetición, las configuraciones, son equiprobables, la probabilidad de ganar es 0,4990, un valor muy próximo al verdadero, 0,4948.

De nuestro trabajo: Juegos de azar, guirguiesca y marlota del Libro de los Dados del Alfonso X el Sabio (Basulto et al, 2007), puede verse la relación que tiene este juego de Marlota con los juegos de Azar de los autores como Pierre Rémond de Montmort (1678-1719) y Abraham de Moivre (1667-1754).

El juego de Riffa [7]

El manuscrito recoge el texto siguiente,

Este es el iuego de la Riffa.

> *"Otra manera de iuego ay: que llaman Riffa que se iuega en esta guisa. El que primero lançare los dados deuelos echar tantas uegadas: fata que lançe par en los dos, desi deue lançar ell otro. Entonce an se de contar los puntos deste dado tercero con los puntos de*

los otros dos dados primeros. E si ell otro que iogare con ell lançando los dados en esta misma guisa echare mas puntos: gana, e si tantos manna, e si menos pierde." (folio 68r)

Libro del ajedrez, dados y tablas, folio 69v

La miniatura muestra a dos jugadores jugando al juego de Riffa, donde uno de ellos está medio desnudo; otros tres jugadores, también medio desnudos, beben agua suministrada por dos mujeres.

Veamos a continuación nuestra interpretación de este juego de Riffa.

En este juego dos jugadores juegan con tres dados de la siguiente forma. El jugador G1 comienza lanzando dos dados hasta que logra sacar la misma puntuación en los dados. A continuación toma otro dado y lo lanza una vez, sumando la puntuación de este último dado con la suma de las puntuaciones de los primeros dados. El jugador G2 lanza, al igual que el jugador G1, primero dos dados hasta sacar par y después lanza un dado, anotando la suma de sus puntuaciones. Gana el juego el que consigue la mayor puntuación. Puede ocurrir que los dos jugadores obtenga la misma suma de puntuaciones, ocurra un empate, con lo que no hay ganador ni perdedor y se deberá comenzar de nuevo el juego (esto no se dice en el texto).

Este juego de Riffa equivale a sacar dos números, con reposición, de una urna que contiene las puntuaciones {2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18}, donde las frecuencias de estas puntuaciones son las recogidas en la Tabla 7-1.

Tabla 7-1

Valores	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Total
Frecuencias	1	1	2	2	3	3	3	3	3	3	3	3	2	2	1	1	36

Ahora, el juego de Riffa consiste en sacar dos números con reposición de la urna tal que si el primer valor es mayor que el segundo, gana el juego G1; si ocurre al revés, gana G2 y si coinciden, empatan. Por cálculo directo encontramos que la probabilidad de ganar el jugador G1 es la misma que la del jugador G2, igual a 0,46451. La probabilidad de que empaten es igual a 0,07098. Si los jugadores cuando empatan no retiran sus apuestas y vuelven a sacar dos nuevos números de la urna hasta que alguno de ellos gane el juego, es fácil de probar que la probabilidad de que el jugador G1 gane el juego es 0,5. Este último resultado explica por qué en este juego no se sortea quién debe comenzar lanzando los dados.

El considerar este juego de Riffa como sacar dos números de una urna cuya composición es la que recogemos en la Tabla 7-1, sirve para ilustrar un muestreo aleatorio simple, con reposición, de tamaño dos. Como sabemos, este tipo de muestreo genera dos variables aleatorias que son independientes e idénticamente distribuidas, y el juego de Riffa pide que calculemos la probabilidad de que una de las variables tome un valor mayor que el valor tomado por la otra.

El juego de Par con As [8]

El manuscrito recoge el texto siguiente,

Este iuego llaman par con as.

> *"Otro iuego a y que llaman par con as e iuegasse desta guisa. El que uenciere la batalla iogara primero. Et si echare par en los dos dados e as en ell otro: gana. Et si no lançare ell otro, desta guisa iogaran; fasta que lance ell uno, e el que primero lo echare: ganara." (Folio 68v)*

Veamos a continuación nuestra interpretación de este juego de Par con As.

Antes de entrar en el juego se debe sortear quién comienza lanzando los tres dados. Vamos a suponer que la "batalla" la ha ganado el jugador G1 y, así, este jugador comienza lanzando los tres dados.

El jugador G1 lanza primero los tres dados, si saca "par" en dos dados y as en el otro, gana el juego. Las configuraciones que hacen ganar este juego son: 1-1-1, 1-2-2, 1-3-3, 1-4-4, 1-5-5 ó 1-6-6. Si G1 no saca una de estas configuraciones, pasa el turno al otro jugador, G2, que lanzará los tres dados buscando sacar una de la configuraciones anteriores. De esta manera se irán turnando hasta que alguno de los jugadores logre una de las configuraciones anteriores, ganando entonces el juego.

Recogemos en la Tabla 8-1 las probabilidades de cada una de las configuraciones de este juego de Par con As.

Tabla 8-1

Configuraciones	Probabilidades de cada Configuración
1-2-2, 1-3-3, 1-4-4, 1-5-5, 1-6-6	3/216
1-1-1	1/216

La probabilidad de que gane el jugador G1, en el caso de que lance el primero, es igual a 0,5192. Por lo tanto, el jugador que lanza primero tiene una ventaja de 0,0192 (1,92%). Cómo en este juego se sortea quién debe comenzar lanzando los dados, si el juego se repite muchas veces, el número de partidas ganadas por G1 se igualará al de partidas ganadas por G2.

Un juego próximo a este Juego de Par con As es el propuesto por el matemático Juan Caramuel (1606-1682). En su obra de Kybeia, que es un tipo de Combinatoria de dados que examina con rigor el juego de dados y los juegos de azar, Caramuel estudia el juego que los españoles llaman Pasa-diez en su Artículo IV (Una traducción de la Kybeia puede verse en la Tesis Doctoral del profesor Camúñez).

Este juego de Pasa-diez consiste en que uno de los jugadores debe lanzar tres dados y el restos de los jugadores pueden hacer apuesta, envidar. Antes de comenzar este juego, los jugadores deben de sortear quién tirará los dados. El juego consiste en lanzar los tres dados y ver si hay dos dados que sacan la misma puntuación; en otro caso se debe volver a lanzar los dados hasta sacar par en dos dados. Si hay dos dados con igual puntuación, se suman todas las puntuaciones de los tres dados, y si la suma supera diez puntos, gana todo lo apostado el jugador que lanza los dados, si es igual o está por debajo de diez pierde, debiendo abonar todo lo apostado. Caramuel prueba que este juego es justo porque la probabilidad de que gane el juego el jugador que lanza los dados es igual a 0,5. Se entiende que Caramuel hace esta afirmación sin considerar que, si se repite el juego muchas veces, es de esperar que cualquier jugador, el que lanza los dados o el resto de los jugadores, gane igual número de partidas porque se sortea quién debe lanzar los dados.

Libro del ajedrez, dados y tablas, folio 69r

En esta miniatura del juego Par con As, vemos que se juega en una casa noble entre señoras y con observadores que son caballeros. Los dados muestran la configuración 3-1-3, que muestra dos dados a la par, con los mismos puntos, 3-3, y el otro dado ha generado el punto 1, es decir, un as.

El juego de Panquist [9]

El manuscrito recoge el texto siguiente,

Este iuego llaman Panquist.

> *"Otra manera ay de iuego que llaman panquist que se iuega en esta guisa: el que uenciere la batalla lançara primero, e ell otro ha de parar quatro paradas una ante: otra, e el que lançare dara la primera suerte all otro, e la segunda tomara pora.*
> *E las suertes que se pueden dar son de siete puntos fata catorce, el que ouiere siete puntos por suerte si echare cinco amas as, o quatro dos e as: leuara las dos primeras, e si echare dos dos e tria leuara las tres. E si echare ternas as: leuara las quatro. E a esta suerte postremera llaman panquist.*
> *Otrossi el que (folio 69r) / ouiere ocho puntos por suerte, si lançare cinco dos e as leuara la primera. E si echare quatro tria e as, leuara las dos. E si lançare seys a amas as, o dos dos e quatro: leuara las tres. E si echare ternas dos leuara las quatro. E a esta suerte postremera llaman panquist.*

> *Otrossi el que ouiere nueue puntos por suerte, si lançare seys dos e as, o cinco tria las: leuara la primera. E si lançare quatro tres e dos: leuara las dos. E si echare dos dos o cinco o ternas alterz: leuara las tres. E si lançare quadernas as, leuara las quatro, e esta suerte postremera llaman panquist.*
> *Otrossi el que ouiere diez puntos e lançare cinco quatro e as o cinco tres e dos: leuara la primera. E si echare seys tria las leuara las dos. E si lançare dos dos e seys, o ternas quatro: leuara las tres. E si echare quadernas dos: leuara las quatro, e a esta suerte postremera llaman panquist.*
> *Otrossi el que ouiere onze puntos, por suerte, e lançare seys tres e dos o cinco quatro e dos: leuara la primera. E si echare seys quatro e as, leuara las dos. E si lançare quinas as o quadernas tria, leuara las tres. E si lançare ternas .v. leuara las quatro, e esta suerte postremera llaman panquist.*
> *Otrossi el que ouiere doze puntos, si lançare seys cinco e as, o seys quatro e dos leuara la primera. E si echare cinco quatro e tria leuara las dos. E si lançare quinas dos o quadernas alterz leuara las tres. E si echare ternas seys leuara las quatro, e a este suerte postremera llaman panquist.*
> *Otrossi el que ouiere treze puntos, e lançare seys cinco e dos: leuara la primera, e si echare seys quatro e tria: leuara las dos. E si lançare senas as, o quinas tria, leuara las tres. E si lançare quadernas cinco leuara las quatro e a este suerte postremera llaman panquist.*
> *Otrossi el que ouiere catorze puntos e lançare seys cinco e tria, leuara las dos. E si echare senas dos o quinas quatro, leuara las tres. E si lançare quadernas seys leuara las quatro, e a esta suerte postremera llaman panquist.*
> *E estas son las suertes por que gana, tan bien el que para como el que lança los dados al que primero uiene su suerte." (Folio 69v)*

Veamos a continuación nuestra interpretación de este juego de Panquist.

En este juego juegan dos jugadores y el que gana la "batalla" debe comenzar lanzando tres dados. Si suponemos que el jugador que lanza es G1, éste deberá lanzarlos hasta obtener una puntuación del conjunto de suertes {7,8,9,10,11,12,13,14}, que se la dará al otro jugador, G2. A continuación, buscara otra <u>puntuación distinta</u> del conjunto suertes que se la quedará para él. En el manuscrito no se habla de que pueda ocurrir "encuentro" como en otros juegos, por lo tanto interpretamos este juego de panquist como que las suertes repartidas son distintas. Una vez repartidas las suertes, el jugador G1 lanzará los tres dados hasta que venga la suerte de G2, que le hará perder, o la suerte de G1, que le hará ganar.

Veamos ahora qué ganan o pierden los jugadores.

En este juego, el jugador G2 propone cuatro apuestas al jugador G1. Estas apuestas están ordenadas de una a cuatro y se denominan: una apuesta, dos apuestas, tres apuestas y las cuatro apuestas. En la Tabla 9-1 recogemos las configuraciones que hacen ganar cada de las apuestas anteriores. Veremos más adelante que estas apuestas están ordenadas de mayor a menor probabilidad de que ocurran. También, de la Tabla 9-1, dentro de cada suerte o puntuación de 7 a 14, las apuestas de mayor tamaño nunca tienen mayor probabilidad que la de menor tamaño. Recordemos aquí que la Bonoloto ofrece unos premios del tipo: acertar tres números, cuatro, cinco o seis números (no considero el complementario), y estos premios van de mayor a menor probabilidad de que ocurran.

Veamos las reglas de este juego de panquist. Si por ejemplo, la suerte del jugador G2 es 7 puntos y la del jugador G1 es 10 puntos, si al lanzar los dados el jugador G1 lanza, por ejemplo, la configuración 3-3-1, entonces habrá ocurrido la puntuación 7=3+3+1 , del jugador

G2, antes que la puntuación 10 del jugador G1, lo que hará ganar el jugador G2. Ahora, como la configuración lanzada por los dados pertenece a la columna de Cuatro primeras Apuestas de la Tabla 9-1, el jugador G2 hará panquist y se llevará estas cuatro apuestas. Si hubiera salido la configuración 6-3-1, entonces ganaría el jugador G1 porque ha llegado antes su puntuación, 6+3+1=10, que la puntuación 7 del jugador G2, La ganancia del jugador G1 será retirar las Dos primera Apuestas porque la configuración 6-3-1 se sitúa en la columna de la Tabla 9-1 con el encabezamiento de Dos primeras Apuestas.

Tabla 9-1

Suertes	Primera Apuesta	Dos primeras Apuestas	Tres primeras Apuestas	Cuatro primeras Apuestas
7		5-1-1, 4-2-1	2-2-3	3-3-1
8	5-2-1	4-3-1	6-1-1, 2-2-4	3-3-2
9	6-2-1, 5-3-1	4-3-2	2-2-5, 3-3-3	4-4-1
10	5-4-1, 5-3-2	6-3-1	2-2-6, 3-3-4	4-4-2
11	6-3-2, 5-4-2	6-4-1	5-5-1, 4-4-3	3-3-5
12	6-5-1, 6-4-2	5-4-3	5-5-2, 4-4-4	3-3-6
13	6-5-2	6-4-3	6-6-1, 5-5-3	4-4-5
14		6-5-3	6-6-2, 5-5-4	4-4-6

Los casos favorables de cada una de las columnas de la Tabla 9-1, según las filas o puntuaciones de las suertes, son los que hemos recogido en la siguiente Tabla 9-2.

Tabla 9-2

Suertes	Primera Apuesta	Dos primera Apuestas	Tres primeras Apuestas	Cuatro primeras Apuestas	Total de Casos Favorables
7		9	3	3	15
8	6	6	6	3	21
9	12	6	4	3	25
10	12	6	6	3	27
11	12	6	6	3	27
12	12	6	4	3	25
13	6	6	6	3	21
14		6	6	3	15
C. F.	60	51	41	24	176

Los casos favorables de la Tabla 9-2 provienen de calcular las permutaciones con repetición que genera cada configuración.

Este Juego de Panquist es igual, en la parte primera donde se reparten las suertes, al Juego de Marlota. En efecto: este juego de Panquist trunca a que las tiradas lancen puntuaciones del conjunto suerte {7,8,9,10,11,12,13,14}, es decir, que si sale otra puntuación se sigue lanzando los dados hasta que venga dos puntuaciones distintas del conjunto suerte, que es igual a la parte primera del Juego de Marlota. En consecuencia, la fórmula [3] del juego de Marlota es aplicable a este juego de Panquist.

Una vez se han repartido las suertes entre los jugadores, el jugador G1 lanza sucesivamente los dados hasta que salga su suerte, que le hace ganar, o la suerte del otro, que la hace perder. Aquí, la fórmula [2], del juego de Triga, nos da la probabilidad que tiene de ganar G1 en el juego de Marlota, condicionada a las suertes repartidas. Ahora bien, lo que necesitamos en el juego de Panquist es la probabilidad de que el jugador G1 gane una cierta apuesta una vez se han repartido las suertes entre los jugadores. Luego, debemos adaptar la fórmula [2] a este nuevo problema tomando en consideración, además de las suertes, las distintas apuestas.

Si S(G_1) y S(G_2) son las suerte distintas de G1 y G2, respectivamente, y el evento de que ocurra la apuesta k, k = 1, 2, 4, lo representamos por A(k), entonces, si se lanzan los dados sucesivamente, debemos calcular la probabilidad de que ocurra antes la suerte S(G_1) que la S(G_2) y además A(k), para un cierto k, y siempre condicionado a las suertes repartidas entre los dos jugadores. La fórmula que proporciona el cálculo es,

$$P_r\left[S(G_1), A(k) \middle| S(G_1), S(G_2)\right] = \frac{CF\left(S(G_1), A(k)\right)}{CF\left(S(G_1)\right) + CF\left(G_2\right)}, \text{ [4]}$$

para k = 1,2,3.

Veamos unos ejemplos, si las suertes son 8 de G1, 9 de G2 y queremos que salga antes la suerte de G1 y además salga la apuesta primera, usando la Tabla 9-2, el valor de la fórmula [4], para k = 1, será,

$$P_r\left[8, A(1) | 8, 9\right] = \frac{6}{21+25} = 0{,}13043.$$

Si las suertes son 11 de G1, 13 de G2 y queremos que salga antes la suerte de G1 y además salga la apuesta segunda, usando la Tabla 9-2, el valor de la fórmula [4], k = 2, será,

$$P_r\left[11, A(2) | 11, 13\right] = \frac{6}{27+21} = 0{,}125.$$

Si las suertes son 12 de G1, 10 de G2 y queremos que salga antes la suerte de G1 y además salga la apuesta segunda, usando la Tabla 9-2, el valor de la fórmula [4], k = 3, será,

$$P_r\left[12, A(3) | 12, 10\right] = \frac{4}{25+27} = 0{,}07692.$$

Por último, si las suertes son 9 de G1, 7 de G2 y queremos que salga antes la suerte de G1 y además salga la apuesta segunda, usando la Tabla 9-2, el valor de la fórmula [4], k = 4, será,

$$P_r\left[9, A(3) | 9, 7\right] = \frac{3}{25+15} = 0{,}075.$$

A partir de la fórmula [2] y la fórmula [4] ya podemos calcular la probabilidad de que el jugador G1 gane cada una de las cuatro apuestas. La Tabla 9-3 recoge las probabilidades de que gane el jugador G1 las distintas apuestas.

Tabla 9-3

Apuestas	Probabilidad de Ganar el jugador G1
I	0,17707721
II	0,14031911
III	0,11440127
IV	0,06663887
Total	0,49843645

Hacemos notar que el total de la Tabla 9-3 corresponde a la probabilidad de que el jugador G1 gane el Juego de Marlota.

De la Tabla 9-3, vemos que las distintas apuestas, ordenadas de menos a más, se corresponden de más a menos en las probabilidades. Como en este juego se sortea quién debe lanzar los dados, se hace batalla, entonces este juego es justo si se juega muchas veces. El jugador G2 que hace de Banca, es decir, propone las apuestas que hemos visto más arriba, tiene una probabilidad de 0,5015 de no perder ninguna de las apuestas. Para la gran apuesta, la IV, la Banca tiene una probabilidad de no perderla igual a 0,933.

Este juego es el Medio Azar [10]

El manuscrito recoge el texto siguiente,

Este iuego llaman medio Azar.

> *"Otra manera ay de iuego de dados que llaman medio azar que se iuega en esta guisa. Los que quisieren iogar an de lançar primeramientre batalla, e el que uenciere lançara primero. E si lançare .xiiii. puntos o dent arriba o siete o dent ayuso, en qual quier manera que uenga cadauna destas suertes sera azar. E de cada azar leuara un tanto, de como pusieren entressi que uala el tanto de un dinero o de un sueldo o un morauedi o dent arriba quanto fuere la postura. E las suertes que son en comedio destas son llamadas suertes, e son estas ocho o nueue e diez e onze e doze e treze. E si por auentura no lançare azar e diere suerte al otro, tomare suerte pora ssi, la que ante uiniere ganara tres tantos. E si desque diere suerte all otro lançare luego azar ante que tome suerte pora ssi, ira de quatro tantos. E si lançare otro azar ira de cinco. E quantos azares lançare uno depos otro, ualdra cadauno un tanto, (folio 70r) / fata que tome suerte por si. E si por auentura ante que tome suerte pora ssi encontrare con la suerte dell otro: lançara de cabo por azar, e lo echare: ganara todos los tantos que y fueren. E si no dar la suerte otra uez, e contara sobre los otros tantos primeros e desta guisa se torna el iuego como de comienço."(Folio 70r)*

Veamos a continuación nuestra interpretación de este juego de Medio Azar.

Es un juego donde se lanzan tres dados. Como en otros juegos que hemos vistos, este juego de Medio Azar también considera el conjunto {3,4,5,6,7, 14,15,16,17,18}, que llamaremos conjunto azar, y el conjunto de {8,9,10,11,12,13} que llamaremos conjunto de suertes, que en el manuscrito dice que están en el medio, es decir, entre la puntuaciones 7 y 14. Si comparamos este conjunto azar con el conjunto azar de otros juegos, como en el Juego de Triga o el Juego de Azar, vemos que en este Juego de Medio Azar, el conjunto azar contiene las puntuaciones 14 y 7, con lo que el conjunto suertes, de este juego de Medio Azar, pierde las puntaciones 14 y 7.

Si el jugador G1 es el que ha ganado la "batalla", cada vez que G1 lance una puntuación del conjunto azar ganará un tanto. Este unidad monetaria será un dinero, un sueldo o un maravedí, aunque podrá ser una cantidad mayor dependiendo de cuál sea la cantidad que los jugadores decidan apostar.

En el caso de que el jugador G1 no lance azar y lance una puntuación del conjunto suerte, ésta será la suerte del jugador G2. Si en el siguiente lanzamiento lanza una puntuación del conjunto suerte, distinta a la dada al jugador G1, ésta será la suerte para G1. Una vez repartidas las suertes, el jugador G1 lanzará sucesivamente los tres dados hasta que venga la suerte de G1, que le hace ganar tres tantos, o la suerte de G2, que hace ganar tres tantos al jugador G2. Y si antes que el jugador G1 logre su suerte, distinta de la del jugador G2, ocurre una puntuación azar, entonces al lanzar los dados sucesivamente, si llega antes la suerte de G1, éste ganará ahora cuatro tantos, que son, los tres tantos de cuando no ocurren azares entre las dos surtes más el tanto debido a que ha ocurrido una puntuación azar. Si entre las dos suertes distintas ocurren dos azares, ahora la posible ganancia de G1 será tres tantos más los dos tantos de los azares, y así sucesivamente.

Vemos entonces que cuando las suertes repartidas son distintas el jugador G1 debe lanzar los dados hasta que surja un ganador en las condiciones indicadas. En el manuscrito, en este caso de suertes distintas, afirma únicamente que "la que ante uiniere ganara", que por el uso que se hace en otros juegos, como en el juego de Triga, Azar, Marlota, Panquist y Guirguiesca, interpretamos que se deben lanzar los dados hasta que venga la suerte de G1, que le hace ganar o la suerte de G2, que hace perder al jugador G1 y, por tanto gana el juego G2.

Si cuando el jugador G1 busca su suerte se encuentra con una suerte igual a la que dio al jugador G2, dos situaciones pueden ocurrir: (1) que el jugador G1 lance los dados y saque una puntuación azar, lo que le hará ganar todos los tantos que allí hubieran y (2) que el jugador lance una puntuación suerte que deberá dársela de nuevo al jugador G2, con lo que el juego comenzará, en parte, de nuevo. En el caso de que el juego comience de nuevo, los tantos generados por las puntuaciones azar, no ganadas, deben acumularse a los posibles tantos que se ganen en esta nueva fase del juego. Esta interpretación que hacemos del juego corresponde a las siguientes líneas:

> *"E si por auentura ante que tome suerte pora ssi encontrare con la suerte dell otro: lançara de cabo por azar, e lo echare: ganara todos los tantos que y fueren. E si no dar la suerte otra uez, e contara sobre los otros tantos primeros e desta guisa se torna el iuego como de comienço."*

En este juego de Medio Azar, el jugador que lanza los dados, G1, gana tantos cuando al comienzo lanza puntuaciones del conjunto azar, hasta que saca una puntuación suerte que se la da al jugador G2. A partir de aquí los tantos se generan porque ocurren azares entre las suertes.

También queremos notar que cuando las suertes han sido repartidas entre los jugadores y son además distintas vemos que a los tantos debidos a los azares se añaden tres tantos, con lo que el jugador que antes saca su suerte se lleva los tantos de los azares y los tres tantos debido a las suerte repartidas. Veremos más adelante, en el juego de Azar Pujado, que ganar tantos por suertes repartidas distintas podrá llevar a ganar un número de tantos que decidan los dos jugadores.

El llamar a este juego Medio Azar se debe a que cuando ocurren puntos de azar en medio de dos suertes repartidas distintas, los jugadores pueden ganar tantos. Y aunque antes de que el jugador G1 dé la suerte al jugador, el jugador G1 puede ganar tantos cuando ocurre una puntuación azar, esto último ya ocurría en otros juegos anteriores, como en el juego de Triga y el juego de Azar. Observamos que el juego puede acabar en las dos siguientes situaciones: (a) cuando las suertes repartidas son distintas; (b) cuando las suertes son iguales y al lanzar los dados el jugador G1 saca una puntuación azar. En cambio, el juego puede comenzar cuando ocurre que las suertes son iguales y además al lanzar los dados G1 saque una puntuación del conjunto suerte. Que comience el juego de nuevo lo hemos visto en el Juego de Azar. Si el juego de Azar debe comenzar cuando ocurre encuentro, en el juego de Medio Azar debe ocurrir además que el jugador G1 al lanzar los dados saque una puntuación del conjunto suerte.

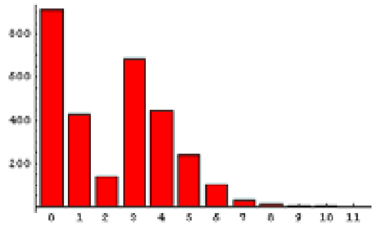
Vamos a calcular la distribución de los tantos que puede ganar el jugador G2 por medio de un programa de ordenador que simula 3000 jugadas de este juego de medio azar. Este cálculo nos permitirá calcular los tantos que espera ganar el jugador G1, que estamos suponiendo que lanza los dados.

La distribución de los tantos ganados por el jugador G1 es la recogida en la Tabla 10-1.

Tabla 10-1

Tantos	0	1	2	3	4	5	6	7	8	9	10	11
Frecuencias	914	424	138	684	443	238	104	31	14	4	5	1
Porcentajes	30,47	14,13	4,6	22,8	14,77	7,93	3,47	1,03	0,47	0,13	0,17	0,03

La gráfica es,



En el gráfico se observar que las frecuencias comienzan con un máximo en cero tantos para descender hasta dos tantos y volver a otro máximo en los tres tantos, esto último es la valoración de tres tantos cuando ocurren dos suertes distintas. Finalmente, las frecuencias descienden hasta cero cuando aumentan los tantos. El número de tantos esperados por el jugador G1 es igual a 2,254.

En este juego de medio azar, al sortear quién debe lanza los dados, cada jugador debe abonar 2,254 tantos para participar en este juego, ya que este valor es el que espera ganar uno de los jugadores al jugar muchas veces.

El juego de Azar Pujado [11]

El manuscrito recoge el texto siguiente,

Este iuego llaman Azar pujado.

"El azar que dizen puiado se iuega desta manera que el medio azar que desuso dixiemos que puian todauia los tantos tan bien por azar, como por qual suerte quiere que uenga la suerte del uno o dell otro. E por esto llaman a este Azar puiado por que ell otro desuso es medio azar." (Folio 70v)

Veamos a continuación nuestra interpretación de este juego de Azar Pujado.

Este juego de Azar Pujado tiene las mismas reglas que el de Medio Azar, sólo se va a diferenciar en el cálculo de los tantos que pueden llevarse los jugadores. Si interpretamos la palabra "pujado" como en su aplicación en las subastas, podemos interpretar este juego de Azar Pujado como que entre los jugadores se producen pujas sobre cuántos tantos se ganará por la puntuación suerte que, en el caso de que las suertes repartidas sean distintas, si al lanzar los dados el jugador G1 llega antes la suerte de G2, entonces éste gana la puja hecha sobre las suertes, mientras que si llega antes la suerte de G1, entonces gana lo pujado el jugador G1. Recordemos que en el juego de Medio Azar, cuando las suertes están repartidas y son distintas, se añaden 3 tantos a los tantos debidos a los azares que pueden venir entre las suertes, lo que podemos interpretar que en este juego de medio azar se puja las suertes por tres tantos. En el juego de Azar Pujado se busca que la puja sea cualquier número de tantos sobre las suertes. Cuando las suertes repartidas son iguales no podemos hacer uso de la pujas.

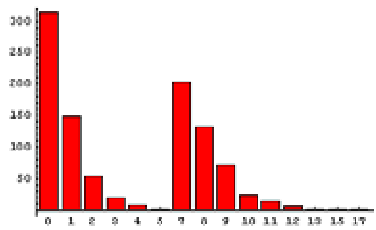
Si suponemos que la cantidad que se puja sobre las suertes es de siete tantos, veamos como se distribuyen los tantos ganados por el jugador G1. Seguimos usando un ordenador para simular 3000 juegos de este juego de Azar Pujado.

La distribución de los tantos ganados por el jugador G1 es la recogida en la Tabla 11-1.

Tabla 11-1

Tantos	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Frecuencias	943	405	151	67	22	7	1	609	401	227	87	51	14	7	6	1	0	1
Porcentajes	31,4	13,5	5,03	2,23	0,73	0,23	0,03	20,3	13,4	7,57	2,9	1,7	0,47	0,23	0,2	0,03	0	0,03

La gráfica es,



Vemos que esta nueva gráfica tiene el mismos comportamiento que la del juego de Medio Azar, variando en que ahora por las suertes distintas se puede ganar siete tantos, de aquí que el segundo máximo ocurra en los siete tantos. El número de tantos esperados por el jugador G1 es igual a 4,119.

En este juego de medio azar, al sortear quién debe lanza los dados, cada jugador debe apostar 4,119 tantos para participar en este juego, así logramos que al jugar muchas veces a este juego, los jugadores esperan ganar los mismos tantos.

Este juego es el Guirguiesca [12]

El manuscrito recoge el texto siguiente,

Este es el iuego que llaman guirguiesca.

> *"Otra manera a y de iuego que llaman guirguiesca que se iuega con dos dados en esta guisa. Los que quisieren iogar an de alançar primeramientre batalla. E el que la uenciere lançara primero, e si lançare senas o seys cinco o la soçobra destos que son dos e as o amas as sera azar. E ganara por el un tanto de qual quantia pusieren entressi que uala. E si por auentura no lançare azar, e echare quatro puntos o cinco, o seys o siete u ocho, o nueue o diez en qual quiere guisa que uengan, cadauna destas sera llamada suerte,e auerla a aquel con qui el iogare, e ell otro parara a ella quanto se quisiere, e si el que lança los dados e echare otra suerte luego a pos ella de tantos puntos como la quel dio esta sera llamada encuentro e leuara lo que y fuere si ouiere otorgado de yr a ello o si sse callare. E si por auentura no lançare encuentro e lançare una delas suertes que desuso dixiemos que eran azares: perder lo a todo. E si non lançare encuentro ni azar e lançare una delas otras suertes aquel la tomara pora ssi, e lançara tantas uezes, fata que uenga la suya o la dell otro, e lançando la suya gana e por la dell otro pierde." (folios 71r y 71v)*

Veamos a continuación nuestra interpretación de este juego de Guirguiesca.

Es este juego de Guirguiesca se lanzan dos dados, las puntuaciones de tipo azar son {2, 3, 12, 11} y las de tipo suertes {4, 5, 6, 7, 8, 9, 10}.

Antes de comenzar este juego de Guirguiesca, los jugadores deben de "hacer batalla" (sortear quién debe lanzar los dos dados). Vamos a llamar G1 al jugador que lanza los dados y G2 al otro. Este juego de Guirguiesca consiste en un primer lanzamiento: (1) Al principio, primer lanzamiento, si sale la puntuación azar, entonces gana el jugador que lanza los dos dados, jugador G1. De esta ganancia sólo se dice que "E ganara por el un tanto de qual quantia pusieren entressi que uala", es decir, una cantidad que hayan acordado ambos jugadores antes de comenza el juego; si G1 lanza una puntuación suerte, en vez de una puntuación azar, entonces debe dársela al otro jugador G2; (2) Si G1 ha lanzado una puntuación suerte en su primer lanzamiento, leemos en el manuscrito que ", e ell otro parara a ella quanto se quisiere", es decir, el jugador G2 hace un envite (apuesta) al jugador G1 donde el valor de esta apuesta lo decide el jugador G2. Este envite del jugador G2 debe ser aceptado por el jugador G1, con lo que G2 obliga al jugador G1 a correr el riesgo de perder la cantidad apostada por G2.

En el segundo lanzamiento: (3) Después del envite de G2, si G1 lanza una puntuación azar entonces pierde lo apostado, antes de comenzar el jugo, y la cantidad del envite de G2. Pero si G1 lanza, en este segundo lanzamiento, una puntuación suerte, entonces pueden ocurrir dos alternativas: (4) Si G1 ha lanzado una suerte que coincide con la suerte del jugador G2, hay encuentro, leemos en el manuscrito que "e si el que lança los dados e echare otra suerte luego a pos ella de tantos puntos como la quel dio esta sera llamada encuentro e leuara lo que y fuere si ouiere otorgado de yr a ello o si sse callare", es decir, ahora el jugador G1 se lleva lo apostado, antes de comenzar el jugo, y la cantidad del envite de G2. Las palabras "si ouiere otorgado de yr a ello o si sse callare" justifican que el envite de G2 obliga al jugador G1 a aceptar el envite independientemente que manifieste su deseo de aceptarlo o que calle; y (5) en el caso de que la surte de G1 sea distinta de la de G2, entonces se deben lanzar los dados hasta que venga la suya, ganando G1, o la de otro, perdiendo G1 y, así, ganando G2.

En la miniatura que recoge este juego de Guirguiesca, folio 71v del Libro del ajedrez, dados y tablas (nuestra página 9), puede verse que el jugador de la derecha ha lanzado los dos dados, nuestro jugador G1, sacando los puntos 1 y 3, que producen la puntuación 4 que sería la suerte para el jugador de la izquierda, nuestro jugador G2. Este último tiene levantada la mano izquierda como haciendo un envite al jugador que ha lanzado los dados.

En la Ley XL, del Ordenamiento de la Tahurerías, se llama a este juego "gargista", diciendo que se trata de un juego de "paradas".

Si nos situamos en el momento donde el jugador G2 tiene su suerte, que vamos a suponer que es de 7 puntos, y hace su envite al otro jugador, G1, sin saber aún qué suerte le llegará a G1, podemos ver en la parte izquierda de la Tabla 12-1 un esquema de esta parte del juego de Guirguiesca.

Tabla 12-1

GUIRGUIESCA		CRAPS	
Puntos	Gana	Puntos	Gana
2	G2	2	Setter
3	G2	3	Seter
4	R	4	R
5	R	5	R
6	R	6	R
7	G1	7	Caster
8	R	8	R
9	R	9	R
10	R	10	R
11	G2	11	Caster
12	G2	12	Setter

De la Tabla 12-1, la parte izquierda, vemos que si el jugador G1 lanza la puntuación 7, ocurre encuentro, entonces gana. Si G1 lanza una puntuación del conjunto {2, 3, 11, 12}, azar, entonces gana G2. Si sale una suerte de los puntos {4, 5, 6, 8, 9}, el jugador G1 tomara para sí su suerte, por ejemplo, 5 puntos, entonces el jugador G1 lanzará los dos dados hasta que venga 7 puntos, ganando G2, o que venga su suerte, 5 puntos, que le hace ganar. Hemos puesto en la celda de la suerte de 5 puntos la letra R que recoge que G1 debe lanzar sucesivamente los dados.

Si consideramos ahora el juego de Craps (Richard Isaac, 1995), un juego donde se lanzan dos dados y juegan dos jugadores llamados Caster y Setter, este último tiene la puntuación 7 como su suerte antes de comenzar el juego.

Cuando sale la puntuación 7 ó 11, llamados cada uno "nick" o natural, gana el jugador Caster; cuando sale los puntos {2, 3, 12}, llamados cada uno "crap", pierde Caster, y así gana Setter; y cuando sale una puntuación suerte del tipo {4, 5, 6, 8, 9}, llamado cada uno "point", por ejemplo, 6 puntos, será la suerte del jugador Caster. A partir de aquí, puede ocurrir que las suertes sean iguales, es decir, que la suerte de Caster sea de 7 puntuaciones como la surte de Setter, que le hace ganar el juego, o que la suerte de Caster sea distintas de 7, por ejemplo, 6 puntos, lo que llevará a que se lancen los dos dados hasta que salga la suerte 7 antes que la 6, ganando Setter, o que venga 6 antes que 7, ganando el juego Caster. Si consideremos todas las suerte de Caster, podemos ver en la parte derecha de la Tabla 12-1 los resultados de este juego de Craps. Se prueba que la probabilidad de ganar Caster este juego de Craps es 0,493, Comparando ambos juegos de la Tabla 12-1, vemos que coinciden en las celdas donde hemos puesto la letra R (repetir los lanzamientos). En otros casos, vemos que el jugador Caster gana si salen las puntuaciones de 7 u 11; mientras que en el juego de Guirguiesca, el jugador G1 gana si sale la puntuación 7. y pierde en la puntuación 11. Por último, las puntuaciones 2, 3 ó 12 que hacen perder al jugador Caster, también hacen perder al jugador G1.

El juego de Craps puede verse en un juego de Hazard de Pierre Rémont de Montmort recogido en las páginas 177-179 de su Essay d'Analyse sur les Jeux de Hazard publicado anónimamente en 1713.

De la Tabla 12-1, donde hemos considerado el juego de Guirguiesca cuando el jugador G2 tiene su suerte igual a 7 puntuaciones, vemos que el juego de Craps es bastante semejante al de Guirguiesca, lo que nos lleva a suponer que el juego de Guirguiesca antecede al juego de Craps. Si ahora consideramos el juego de Guirguiesca cuando el jugador G2 puede tener algunas de las suertes {4, 5, 6, 7, 8, 9, 10} y propone un envite al jugador G1, entonces podemos representar los posibles resultados de esta parte del juego mediante la siguiente Tabla 12-2.

Tabla 12-2

	G2						
G1	4	5	6	7	8	9	10
2	G2	G2	G2	G2	G2	G2	G2
3	G2	G2	G2	G2	G2	G2	G2
4	G1	R	R	R	R	R	R
5	R	G1	R	R	R	R	R
6	R	R	G1	R	R	R	R
7	R	R	R	G1	R	R	R
8	R	R	R	R	G1	R	R
9	R	R	R	R	R	G1	R
10	R	R	R	R	R	R	G1
11	G2	G2	G2	G2	G2	G2	G2
12	G2	G2	G2	G2	G2	G2	G2

En esta Tabla 12-2, recogemos las posibles suertes de G2, en las columnas, y para cada suerte, por ejemplo, 5 puntos del jugador G2, tenemos que si el jugador G1 lanza una puntación azar, entonces gana el jugador G2, si saca una puntuación suerte igual a 5 puntos, gana G1 y, finalmente, si saca una puntuación suerte distinta de 5 puntos, por ejemplo, 7 puntos, entonces G1 debe lanzar los dos dados sucesivamente hasta que lance 7 puntos, que le hace ganar, o 5 puntos que le hace perder (gana G2). Esta repetición lo indicamos en la Tabla 12-2 con la letra R de "repetición".

Hemos calculado, para cada suerte del jugador G2, la probabilidad de que el jugador G1 gane el envite del jugador G2. Estas probabilidades son recogidas en la siguiente Tabla 12-3.

Tabla 12-3

Suertes de G2	4	5	6	7
Probabilidad de Ganar G1	0,5371	0,4924	0,461	0,4374
Suertes de G2	10	9	8	

De la Tabla 12-3, observamos que las Suertes {5, 9, 6, 8, 7} de G2 le dan ventaja, mientras que para el resto de las Suertes de G2 las ventajas son para el jugador G1. También de la Tabla 12-3 vemos que cuando las suertes de G2 son 5 ó 9, entonces en envite de G2 es bastante justo para ambas jugadores. Recordemos que en el juego de Craps la probabilidad de ganar que tiene el jugador Caster es 0,493, que está muy cercana a la probabilidad 0,4924 de ganar G1 cuando las suertes de G2 son 4 ó 9.

La probabilidad de que el jugador G2 gane su envite cuando su suerte es S(G2) puede se calculada por la fórmula siguiente,

$$\frac{CF(Azar)}{36}+\sum_{k=4;k\neq S(G2)}^{10}\frac{CF\big(S(G2)\big)*CF\big(S(k)\big)}{36*(CF\big(S(G2)\big)+CF\big(S(k)\big))},$$

donde S(k) son las surtes del conjunto {4, 5, 6, 7, 8, 9, 10}, que puede tener el jugador G2, y los casos favorables del conjunto Azar son 6.

A falta de conocer las apuestas, tanto al principio del juego como cuando el jugador G2 hace su envite, nos podemos interesar por la probabilidad de que el jugador G1 gane este juego de Guirguiesca, es decir, que G1 saque una puntuación azar en su primer lanzamiento o que saque un suerte igual a la de G2 en su segundo lanzamiento o que saque una suerte distinta a la de G2 en su segundo lanzamiento que le lleve a ganar. Si calculamos estas tres probabilidades y las sumamos obtenemos una probabilidad de 0,566 de que G1 gane este juego de Guirguiesca. Al superar esta probabilidad el valor 0,5, muestra que este juego de Guirguiesca es ventajoso para el jugador que lanza los dados. Esta ventaja explica por qué, antes de comenzar el juego, los jugadores deben de sortear quién debe lanzar los dos dados, lo que, a la larga, hará que los jugadores ganarán el mismo número de partidas. Este último cálculo nos permite afirmar que el juego de Guirguiesca es semejante al de Triga cuando jugamos con dos dados, siempre que se elimine la "batalla", no sea de interés las cantidades que los jugadores pueden ganar y además no influya el orden en que se reparten las suertes.

El Libro de los Dados finaliza con la frase siguiente.

"En estos .xii. iuegos delos dados que aqui auemos puesto, se pueden entender todos los otros que iuegan en las otras tierras que son fechos o se pueden fazer daqui adelant de que nos non sabemos." (Folio 71v)

Bibliografía

BASULTO, J.; J. A. CAMÚÑEZ Y J. ORTEGA (2006). "Azar Game in the Book of Dice of Alfonso X The Learned. Its relation with the Hazard Games of Montmort, Cotton, Hoyle, De Moivre and J. Bernoulli". Math. & Sci. Hum.-Mathematics and Social Sciences. No. 174, 5-24.

BASULTO, J.; J. A. CAMÚÑEZ Y J. ORTEGA. (2007). "Juegos de azar, guirguiesca y marlota del Libro de los Dados del Alfonso X El Sabio". Alcanate. Revista de Estudios Alfonsíes. 89-116.

BELLHOUSE, D. R. (2004). "Decoding Cardano's Liber de Ludo Aleae". Historia Matemática. 1-22.

BELLHOUSE, D. R. (2000). "De Vetula: a Medieval Manuscript Containing Probability Calculations". International Statistical Review. 68, 123-136.

BERNOULLI, J. (1713). The Art of Conjecturing. Together with Letter to a Friend on Sets in Court Tennis. Translated with an introduction and notes by Edith Duddley Sylla. The Johns Hopking University Press. 2006.

CARDANO, G. (1564). *Liber de Ludo Aleae*. First printed in Opera Omnia, vol. 1, 1663. Translated into inglish by S.H. Gould in Ore (1953), reprinted in The book on game of chance, 1961, Holt, Rinehart and Winston, New Tork.

COTTON, C. (1674). *The compleat gamester: or, instructions how to play at billads, trucks, bowl,...and chess. Together with all manner pf usual and most gentile games either on cards or dice.* London, print by A.M. for R. Cutler.

HUYGENS, C. (1657). *De Ratiociniis in Ludo Aleae*, printed in Exercitationum Mathematicarum by F. Van Schooten, Elsevirii, Leiden.

MOIVRE, A, DE (1718). *The Doctrine of Chances or, A Method of Calculating the Probability of Events in Play*. Pearson. London.

MONTMORT, P.T. DE (1714). *Essay d'Analyse sur les Jeux de Hazard*. Second Edition.Published anonymously. Reprinted by Chelsea, New York, 1980.

MORA CHARLES, M. S. DE (1989). *Los inicios de la Teoría de la Probabilidad: siglos XVI y XVII.* Servicio Editorial de la Universidad del País Vasco, Bilbao.

MORA CHARLES, M. S. DE (1992). "Quelques jeux de hazard selon Leibniz ». Historia Mathematica 19, 125-157.

RICHARD ISAAC. (1995). *The Pleasures of Probability*. Spinger Verlag.

Capítulo 13

Lo que la astronomía regaló a la estadística

GABRIEL RUIZ GARZÓN
Universidad de Cádiz

Antecedentes

En la Historia de la Estadística y la Probabilidad son varios los científicos famosos que aportaron su saber tanto a la Astronomía como a la Estadística.

Así por ejemplo, el sabio holandés Christiaan Huygens (1629-1695), que entre los estadísticos es famoso por su monografía de veinte páginas en latín titulada *"De ratiociniis in ludo alae"*, donde aparece el concepto de *"expectatio"*, o valor esperado de un juego, que después se tradujo por el de esperanza matemática como una generalización del concepto de media aritmética, sin lugar a dudas, el concepto más importante de la Estadística.

Como astrónomo, Huygens descubrió en 1655 el primer satélite de Saturno y estableció la verdadera forma de sus anillos. También descubrió la nebulosa Orión, siendo el primero en indicar que las estrellas son otros soles.

Quizás Edmund Halley (1656-1742), nombrado Astrónomo Real en 1721 y Director del Observatorio de Greenwich, es más conocido por el público en general por descubrir el cometa que lleva su nombre que por sus aportaciones a al Estadística. En 1682 se dio cuenta que el cometa de ese año era el mismo que el de 1607 y predijo que aparecería 76 años después, como así fue. La última vez que ha aparecido hasta la fecha fue en 1986 y la próxima vez que nos visite será en el año 2061.

También propuso la utilización de tablas de la órbita de la Luna para determinar el cálculo de la longitud en el mar o el uso del tránsito de Venus por el Sol para determinar la distancia del Sol a la Tierra, mediante cálculos trigonométricos. El tránsito de Venus es el paso de Venus por delante del Sol, visto desde la Tierra

Como estadístico, Edmund Halley, ayudado por Leibniz, obtuvo los datos de mortalidad de la ciudad alemana de Breslau, a fin de calcular el *valor de las pensiones* individuales y mancomunadas en función de la edad de los adquirientes. Su labor le llevó a calcular la primera tabla de mortalidad en un artículo titulado *"An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslau"*, publicado en la *Philosophical Transactions* del año 1693 y que se haya depositado en la Biblioteca del Real Observatorio de S. Fernando.

Gauss, Legendre, el método de mínimos cuadrados, los cometas y los asteroides

Carl Friedrich Gauss (1777-1855), es llamado el *"Príncipe de las Matemáticas"*. Gauss muere en Göttinguen en 1855 ciudad de la que fue director de su Observatorio Astronómico.

Otro matemático Adrien-Marie Legendre (1752-1833), mantuvo con Gauss una agria disputa sobre la prioridad en el descubrimiento del método de mínimos cuadrados, ver (García, 2002). Legendre fue el primero en publicar el método de mínimos cuadrados en 1805, en un apéndice del libro sobre la órbita de los cometas: *"Nouvelles méthodes pour la détermination des orbites des cométes"*.

El método de mínimos cuadrados era en sus comienzos una técnica geodésica. Era una técnica para resolver sistemas de ecuaciones donde el número de ecuaciones superaba al de incógnitas, algo que ocurría con bastante asiduidad en Astronomía. Varios eran los problemas que desembocan en dichos sistemas de ecuaciones indeterminados.

Uno de ellos es el problema de la figura de la Tierra, y ligado a él la medición del arco del meridiano terrestre, y a su vez relacionado con éste último, el problema de la introducción del Sistema Métrico Decimal. La cuestión era si la forma del planeta era achatada en los polos o en el ecuador.

Para solucionar este problema Laplace en 1780 y Legendre en 1805, utilizaron la siguiente ecuación:

$$a = x + y\,\mathrm{sen}^2 L \qquad [1]$$

donde $a = S/d$ es la longitud del arco en módulos por grado de latitud, x= la longitud de un grado en el ecuador e y= el exceso de un grado en el polo sobre uno del ecuador. Se trataría de encontrar las constantes x e y.

Con los cálculos de dos científicos franceses, Delambre y Méchain, hemos elaborado la siguiente tabla que nos da la longitud S de los cuatro segmentos consecutivos del arco de meridiano que atraviesa París, medidos en módulos (1 módulo $\cong$ 12'78 pies), los grados de latitud d y la latitud del punto medio L de cada segmento de arco:

	MÓDULOS S	***GRADOS* d**	PUNTO MEDIO L
De Dunquerque a Panteón	62472'59	2'18910	49º 56' 30"
De Pantheon(París) a Evaux	76145'74	2'66868	47º 30' 46"
De Evaux a Carcassone	84424'55	2'96336	44º 41' 48"
De Carcassone a Barcelona	52749'48	1'85266	42º 17' 20"

Con los resultados de las observaciones de Delambre y Méchain que figuran en la tabla anterior, substituidas en la ecuación anterior [1], dan las llamadas cuatro *ecuaciones de condición*:

$$\begin{aligned} x + y(0'585821) - 28538'02476 = 0 \\ x + y(0'543800) - 28533'11000 = 0 \\ x + y(0'494705) - 28489'46804 = 0 \\ x + y(0'452752) - 28472'29389 = 0 \end{aligned},$$

Si aplicamos el método de mínimo cuadrados, esas 4 ecuaciones se pueden substituir por las siguientes dos ecuaciones del *sistema de ecuaciones normales*:

$$\begin{aligned} 4x + y(2'077078) - 114032'8967 = 0 \\ 2'077078x + y(1'088621) - 59219'24971 = 0 \end{aligned}$$

cuyas soluciones son

$$x = 28227'13109, \quad y = 541'3241$$

y así

$$a = 28227'13109 + 541'3241 \operatorname{sen}^2 L .$$

Por tanto, gracias al método de mínimos cuadrados, un sistema formado por cuatro ecuaciones con dos incógnitas las hemos transformado en un sistema de dos ecuaciones con dos incógnitas de fácil solución.

Otro problema que conlleva la resolución de un sistema de ecuaciones como el anterior consiste en determinar matemáticamente y con mucha exactitud los movimientos de la luna. Esto es debido a que la luna se ve afectada de un pequeño balanceo en torno a su eje, llamado *libración*, que se percibe desde la Tierra. Esto hace que la luna no presente exactamente el mismo hemisferio hacia la Tierra.

El objetivo de conocer exactamente el movimiento de la Luna era poder describir posiciones futuras de la luna y poder utilizar dichas tablas o almanaques para poder determinar la longitud en la navegación marítima.

Esto le suponía resolver al astrónomo Johann Tobias Mayer (1723-1762), veintisiete ecuaciones con sólo 3 incógnitas, ecuaciones resultantes del mismo número de observaciones del cráter Manilius efectuadas desde el 11 de Abril de 1748 hasta el 4 de Marzo de 1749. El método de resolución consistía en agruparlas en 3 grupos de 9 ecuaciones cada uno, sumar las ecuaciones de cada uno de los tres grupos y resolver las 3 ecuaciones resultantes. Con el método de mínimos cuadrados el procedimiento se simplificaba y no se dejaba a la buena o mala mano que tuviera el matemático de turno en la combinación de ecuaciones hasta llegar a un sistema compatible determinado.

En 1749, el matemático Leonard Euler ganó un premio de la Academia de París en que intentaba dar explicación a movimientos de aceleración y retardo de los planetas Júpiter y Saturno en el transcurso del tiempo. Cada año había que recalcular nuevas tablas porque los planetas se desviaban de las órbitas elípticas marcadas. Dicho retraso o adelanto es debido a la interacción de las fuerzas gravitatorias de ambos planetas y del Sol. El problema de describir las órbitas de estos planetas es conocido como el *Problema de los Tres Cuerpos* (Júpiter, Saturno y el Sol). La posible inestabilidad de las órbitas de estos cuerpos podría provocar el choque entre ellos.

Esto suponía a Euler resolver un sistema formado por 75 ecuaciones con 8 incógnitas mediante un sistema parecido al utilizado por Mayer de combinación de ecuaciones. Esto sólo estaba al alcance de un genio como Euler. El método de mínimos cuadrados reducía el ingente trabajo operatorio que de otra manera convertía al problema en casi irresoluble.

Otro problema relacionado con la aparición del método de mínimos cuadrados fue el de las órbitas descritas por los asteroides.

En 1801, el astrónomo italiano Joseph Piazzi (1746-1826), descubrió el asteroide al que le pusieron el nombre de Ceres. El 28 de Marzo de 1802, el médico alemán W.H. Olbers (1758-1840) descubrió el siguiente asteroide, al que le pusieron por nombre Pallas. La mayoría de los asteroides orbita entre Marte y Júpiter. Son objetos rocosos y metálicos que se forman desde los inicios del Sistema Solar.

A causa de la situación desfavorable del asteroide Ceres respecto del Sol, sólo pudieron practicarse las observaciones durante 40 días.

Al haberlo perdido tan pronto, los astrónomos se enfrentaron al problema de calcular sus posiciones a partir de pocas observaciones. Gauss calculó por mínimos cuadrados la trayectoria de Ceres de tal manera que cuando el asteroide apareció por el otro lado del Sol, los astrónomos lo encontraron cuándo y dónde Gauss les había dicho. Para el cálculo de la órbita de Pallas, Gauss también planteó un sistema inicial de 12 ecuaciones con 6 incógnitas. Pero utilizando el método de mínimos cuadrados, se resolvía este problema donde el número de ecuaciones era mayor que el de incógnitas. La Estadística venía a resolver problemas de Astronomía.

Daniel Bernoulli, Laplace, los contrastes de significación y la inclinación de la órbita de los planetas

La Tierra, al desplazarse en torno al Sol, se mantiene dentro de un plano, llamado *eclíptica*. La razón de ese nombre es que los eclipses de Sol o de Luna sólo se producen cuando ésta atraviesa, en su órbita, la eclíptica, pues sólo entonces puede pasar entre la Tierra y el Sol y eclipsar éste último, o directamente por detrás de la Tierra de forma que quede dentro de la sombra de ésta y siendo por tanto eclipsada.

¿Estarían en la eclíptica las órbitas de los demás planetas? La respuesta es: no. La órbita de cada planeta tiene su propio plano independiente, que no se parece al de la órbita de ningún otro. El plano de cada planeta corta la eclíptica según un cierto ángulo, como vemos en la siguiente tabla con los planetas conocidos en tiempos de Daniel Bernoulli (1700-1782):

Planeta	*Inclinación respecto a la eclíptica (grados)*
Mercurio	6º 54’
Venus	3º 22’
Tierra	-
Marte	1º 50’
Júpiter	1º 20’
Saturno	2º 32’

Esto demuestra que aunque el sistema solar no sea un objeto perfectamente plano, está muy cerca de ello.

Imaginemos un planeta que sigue una órbita cuya inclinación con respecto a la eclíptica es un ángulo pequeño. Durante la mitad de su órbita se está moviendo por encima de la eclíptica. Luego, en un punto de la órbita, atraviesa la eclíptica para situarse por debajo de ella durante otra mitad de la órbita, volviendo luego a situarse por encima, y así sucesivamente. Los dos lugares donde la órbita del planeta atraviesa a la eclíptica se llaman *nodos*, y están situados en puntos opuestos de la órbita. Pero, a qué causa es debida tal inclinación de las órbitas.

Daniel Bernoulli miembro de la excelsa familia matemática de los Bernoulli, ganó diversos premios de la Academia de París sobre temas tan diversos como: la forma de los barcos, ensayos sobre el magnetismo, sobre las mareas, etc.

Concretamente en 1734, Daniel Bernoulli gana junto con su padre, el premio establecido por la Academia de Ciencias de París para responder a la pregunta de cuál es la causa física de la inclinación de los planos de las órbitas de los planetas.

La solución de Daniel queda explicada en un artículo primero escrito en latín y después traducido al francés. Los razonamientos de Daniel están ligados a los inicios de los contrastes de significación, como ahora pasamos a ver.

Su objetivo se centra en contrastar la siguiente hipótesis estadística:

- H_0 : La inclinación de las órbitas de los planetas respecto de la eclíptica es debida al azar, es decir, existe igual facilidad de las inclinaciones de las órbitas.
- H_1 : Tal inclinación es debida a una ley física, existe una causa primitiva que influye en las observaciones.

Para tratar este problema, Daniel Bernoulli, de entre todas las órbitas planetarias, busca las dos que se cortan bajo un ángulo más grande y después calcula la probabilidad de que las otras órbitas estén contenidas al azar entre estos dos límites.

De este modo, encuentra que el ángulo más grande entre los planos orbitales se da entre Mercurio y la Tierra con 6º 54’. Consideradas las órbitas planetarias colocadas en la esfera celestial al azar, la probabilidad de que una órbita planetaria caiga dentro de una zona de anchura 6º 54’ es:

$$p = \frac{\textit{área del segmento esférico}}{\textit{área de la esfera}} = \frac{2\pi r h}{4\pi r^2} = \frac{h}{2r} = \frac{1}{2}\, sen\ 6^{\circ}\ 54' \cong \frac{1}{17}.$$

Por lo tanto, la probabilidad de que las órbitas de esos cinco planetas se desvíen como mucho 6º 54’ de la eclíptica es

$$p(\textit{órbitas 5 planetas} \leq 6^{0}54' | H_0) = \frac{1}{17^5}.$$

Daniel, en la traducción al francés del artículo añade que es imposible determinar con tal exactitud tal probabilidad porque los movimientos de los nodos cambian los límites de las órbitas. No obstante, si todos los nodos se mantuvieran estacionarios en un punto, todos los planos se intersecarían en una línea común, entonces el problema se podría simplificar, ya que se podrían considerar los ángulos distribuidos al azar entre 0º y 90º, y por tanto, la probabilidad de que una órbita planetaria caiga dentro de una zona de anchura 6º 54’ es:

$$p = \frac{6^{\circ}\,54'}{90^{\circ}} \cong \frac{1}{13}.$$

Luego, la probabilidad de que las órbitas de esos cinco planetas se desvíen como mucho 6º 54' de la eclíptica es

$$p(\acute{o}rbitas\ 5\ planetas \leq 6^{0}54' | H_0) = \frac{1}{13^5}.$$

Daniel Bernoulli piensa que quizás la probabilidad real se encuentre entre $1/17^5$ y $1/13^5$. De cualquier modo se debe rechazar la hipótesis nula de que el fenómeno es debido al azar y busca la causa física de las inclinaciones de las órbitas de los planetas en los efectos de la gravitación y de la atmósfera solar, que provocarán que a la larga todas las órbitas planetarias coincidan con el plano solar.

Pierre Simón Laplace (1749-1827), fue maestro y ministro de Napoleón. En su *"Mecánique Celéste"* aplica las herramientas matemáticas a la Física del sistema solar. Sobre esta obra, el General Bonaparte escribía a Laplace: *"Los primeros seis meses libres que pueda disponer, serán empleados en leer vuestra bella obra"*. Al acabar su lectura, Napoleón le preguntó a Laplace por qué Dios no aparecía en dicho trabajo de Astronomía, a lo que el sabio le respondió: *"No he tenido necesidad de dicha hipótesis"*.

Laplace en un artículo publicado en 1776 *"Mémoire sur l'inclinaison moyenne des orbites des comètes"*, también se ocupa de dar una explicación a los diferentes grados de inclinación de los planos de las órbitas de los planetas respecto de la eclíptica.

Los argumentos de Laplace también se hallan en los orígenes de los contrastes de significación. Concretamente, utilizará un contraste de significación pero para la media. Supuesto que la media observada es más pequeña que el valor esperado $\bar{x}$ bajo la hipótesis nula, el test de Laplace consiste en calcular la probabilidad de que se dé una desviación del valor esperado tan grande como el observado. Si esta probabilidad es pequeña se rechaza la hipótesis nula basándonos en que lo observado resulta improbable bajo esa hipótesis.

De los datos de la tabla que D. Bernoulli utilizó, obtenemos una inclinación media de 3º.192 y Laplace calcula por tanto

$$p\left(\bar{x} \leq 3^{\circ}.192 \middle| H_0\right) = \frac{1}{684500},$$

que es más pequeña que $1/13^5$, dato éste último aportado por Daniel Bernoulli, pero que también nos llevaría a rechazar la hipótesis nula. Luego la inclinación de las órbitas de los planetas no es debida al azar, existiendo una causa que la origina. Una vez más un problema astronómico está en el inicio de los contrastes de significación.

Michell y la probabilidad asociada a la aparición de estrellas dobles

John Michell (1724-1793) fue clérigo de Thornhill, villa inglesa en las cercanías de Leeds. Fue el primer hombre que predijo la existencia de *agujeros negros*. Planteó la idea de una estrella invisible en un artículo publicado en 1784 en la *Philosophical Transactions* titulado *"On the means of discovering the distances…"*.

Otro tema de estudio de Michell fueron las *estrellas dobles.* En 1767 publicó un artículo en la *Philosophical Transactions* titulado *"An Inquiry into the probable parallax…"* donde trataba temas relacionados con las estrellas dobles.

Éstas habían sido vistas antes por otros astrónomos, cuando al enfocar sus anteojos a ciertas estrellas, habían percibido dos imágenes donde se creía que debía existir una sola. Se trataba de pares de estrellas.

La opinión de la época postulaba que se trataba de dos estrellas muy alejadas en el espacio una de la otra, pero que un efecto de perspectiva de nuestra vista las hacía aparecer próximas. Otros, en cambio, opinaban que ambas estrellas eran, en realidad, un sistema formado por dos cuerpos que giraban uno en torno del otro ligados por una misma atracción.

Se trataría, por tanto, según Michell, de contrastar la siguiente hipótesis estadística:

H_0 : Las estrellas se encuentran dispersas al "azar" (es decir, de manera uniforme en el cielo), esto se manifestaría en que cualquier estrella tendría la misma probabilidad de estar en un sitio como en otro del cielo. Las estrellas dobles serían debidas a gran variedad de leyes estelares o de posición. Las estrellas dobles sólo serían dobles ópticamente, es decir, parecerían dobles aquellas que aparecen juntas cuando son observadas desde la Tierra pero no serían dobles físicamente, sólo se encontrarían en la misma dirección.

H_1 : La distribución de las estrellas es debida a una ley física, existe una causa primitiva que influye en las observaciones y que hace que las estrellas se agrupen en algunas zonas formando grupos, las estrellas que son dobles ópticamente lo son dobles realmente.

Para ello, John Michell calculará cómo de probable es descubrir dos estrellas a tan corta distancia aparente como parece que sucede.

Asume que la probabilidad de que cualquier estrella se halle situada dentro de un área de la esfera celeste es proporcional a esta área.

Sea r el radio de una pequeña área circular sobre la superficie de una esfera de radio unidad. Ignorando la curvatura de una pequeña área, la probabilidad, bajo la hipótesis nula, de que una estrella esté situada dentro de una pequeña área circular de radio r radianes es

$$p_r = \frac{\text{área circular}}{\text{área de la esfera}} = \frac{\pi r^2}{4\pi 1^2} = \left(\frac{r}{2}\right)^2 .$$

Si expresamos r en vez de en radianes en minutos, la probabilidad de que una estrella esté situada dentro de una pequeña área circular de radio x minutos bajo la hipótesis nula es

$$p_x = \frac{\pi r^2}{4\pi 1^2} = \left(\frac{r}{2}\right)^2 = \left(\frac{x\pi}{2\times 60\times 180}\right)^2 = \left(\frac{x}{6875.5}\right)^2 .$$

Por tanto, la probabilidad de que cualquier estrella A diste de cualquier otra estrella B menos de un grado se consigue sustituyendo en la fórmula anterior $x=60'=1^o$ y es igual a

$$p_x = p\left(d(A,B)\le 1^0 \middle| H_0\right) = \frac{\pi r^2}{4\pi 1^2} = \left(\frac{x\pi}{2\times 60\times 180}\right)^2 = \frac{1}{13131} .$$

Así pues, la probabilidad de que de entre n estrellas no haya ninguna que diste menos de un grado de otra dada es $(1-p_x)^n$. Dado que cualquiera de esas n estrellas puede ser tomada como referencia de las demás, Michell concluye que la probabilidad en cuestión es $(1-p_x)^{n\times n} = q_x^{n\times n}$.

Aplica la citada fórmula a β Capricorni, una doble estrella dentro de una clase de 230 estrellas del mismo brillo y que en la actualidad puede verse separada con binoculares. Ya que la distancia entre las dos estrellas es como mucho de $x = 3.\widehat{3}'$, entonces $p_x = 1/4254525$, y por tanto, $q_x^{230\times 230} = 80/81$ es la probabilidad de que ninguna de las dos estrellas se encuentre dentro de esa pequeña distancia de cualquier otra. Así pues, la probabilidad bajo la hipótesis nula, de observar un grupo como β Capricorni es de sólo 1/81.

$$p\left(d(A,B) \leq 3.\widehat{3}' \middle| H_0\right) = \frac{1}{81}.$$

John Michell escribe que no es posible mediante una observación directa decidir si dos estrellas observadas se encuentran tan juntas para ser sólo una doble estrella óptica o una doble estrella real, esto es, si un par de estrellas están lo suficientemente cercanas para que el movimiento de una influya en la otra. Pero es altamente probable que alguna doble o múltiple estrella sí lo sea físicamente.

La conclusión era que las estrellas no estaban distribuidas al azar. Existe alguna ley que tiende a producir las proximidades observadas, y que las proximidades actuales no son meramente ópticas y aparentes. Las estrellas que parecían muy cercanas una de la otra debían estarlo en realidad en casi todos los casos. Sólo en unos pocos casos excepcionales podía darse el accidente de que sólo estuviesen en la misma dirección.

Esas estrellas que están realmente muy cerca una de la otra son llamadas *estrellas binarias.*

El interés por las estrellas dobles tiene mucho que ver con la medición de la paralaje estelar. La *paralaje* es el cambio aparente de la posición de un objeto próximo en comparación con un objeto más distante, cuando el observador cambia la posición desde la que observa el objeto. Se puede optar por observar dos estrellas que estuvieran muy juntas. Era lógico suponer que no estaban muy cerca una de la otra, sino que sencillamente se encontraban en la misma dirección con respecto a la Tierra. Una de las dos, la más oscura, estaría tan lejos que no mostraría prácticamente ninguna paralaje. Por tanto, se la podría considerar estacionaria. La más brillante del par, en cambio, estaría bastante cerca para mostrar una paralaje detectable. Así, presentaría un pequeño desplazamiento anual con referencia a la estrella oscura estacionaria cercana a ella. Por el tamaño de la paralaje es posible determinar, mediante cálculos trigonométricos, la distancia entre nuestro planeta y una estrella, por ejemplo.

Luego la probabilidad también se aplicó a la aparición de las estrellas dobles.

Otros protagonistas

Otro protagonista de la estrecha relación entre Astronomía y Estadística fue el astrónomo William Herschel (1738-1822). Nacido en Hannover, una región alemana que estaba bajo el domino del Rey Jorge II de Gran Bretaña. Estudió música y en 1757 huyó a Inglaterra a fin de no ser reclutado en el ejército para combatir a los franceses. William Herschel se hizo sus propios telescopios porque no tenía dinero para comprarse uno bueno y acabó construyendo el mejor de su tiempo. Por cierto, que William Herschel construyó el primer telescopio para el Observatorio Astronómico de Madrid.

También, William Herschel descubrió un nuevo planeta, al que quiso poner el nombre del monarca inglés *"Estrella de Jorge"*, pero al que se acabó llamando *Urano* a propuesta del astrónomo Johann Elert Bode (1747-1826), ya que según la mitología griega, Urano era el padre de Crono (Saturno). Al igual que Saturno era el padre de Zeus (Júpiter) y éste a su vez era el padre de Ares (Marte), Afrodita (Venus) y Hermes (Mercurio), dioses que daban nombre a los planetas interiores.

Un motivo de estudio de William Herschel fue la forma de la nuestra galaxia. A primera vista, parecería que las estrellas se esparcían al azar por el espacio infinito en todas las direcciones. Podemos ver estrellas en todas las direcciones, y con el telescopio se pueden ver más estrellas en todas las direcciones. Sin embargo, Herschel presuponía que la forma global de nuestro sistema sería la de una piedra de afilar o una lente. Si desde la Tierra miramos a lo largo del eje corto de la lente, veríamos relativamente pocas estrellas. Si miramos en la dirección del eje largo, veríamos un enorme número de estrellas.

Para verificar dicha hipótesis, tratar de contar todas las estrellas en todas las direcciones era imposible. De modo que en 1784 efectúo un *muestreo* del cielo. Eligió 683 regiones, dispersas por todo el cielo, y contó las estrellas visibles con su telescopio en cada región. Halló un número de estrellas por unidad de superficie mayor en el plano de la Vía Láctea y un menor número de estrellas en ángulos rectos a este plano. Es decir, gracias a razonamientos inferenciales, William Herschel dio forma a nuestra galaxia.

También es reseñable la labor de Christian Kramp (1760-1826). En 1809 fue nombrado profesor de Matemáticas en Estrasburgo, ciudad en la que nació y murió. Sus trabajos en Matemáticas se centraron en estudiar la función factorial generalizada. Fue uno de los primeros en utilizar la notación *n!* para el factorial de un número.

Para la Estadística tiene importancia su manual para el cálculo de refracciones astronómicas donde aparecen tabulados los valores de la integral $\int_t^\infty e^{-t^2}\,dt$. Esta tabla será utilizada por otros estadísticos como Cournot para poder tabular la distribución normal.

En las tablas de refracciones astronómicas aparecen las correcciones que deben efectuarse a los desplazamientos aparentes llevados a cabo por causa de la refracción, en todas las altitudes o en cada situación en que esté situado el cuerpo solar observado en cuestión.

La refracción astronómica es la modificación de la dirección de una estrella ocasionada por la desviación del rayo luminoso que va de la estrella a la Tierra al atravesar la atmósfera terrestre. Este fenómeno hace que el Sol y las estrellas se vean siempre por encima de su posición real. Luego los astrónomos ayudaron a tabular la Normal.

Bibliografía

BERNOULLI, D. (1735): "Recherches physiques et astronomiques sur le problem propose pour la seconde fois par l'Académie Royale des Sciences de Paris: Quelle est la cause physique de l'inclinaison des plans des orbites des planetes par rapport au plan de l'equateur de la révolution du soleil autour de son axe; Et d'où vient que les inclinaisons de ces orbites sont diferentes entre elles", *Recuel des pièces qui ont remporté les prix de l'Acadèmie Royale des Sciences,* 3, 303-326.

EULER, L. (1749): "Recherches sur la question des inégalités du mouvement de Saturne et Júpiter, sujet propose pour le prix de l'anné 1748, par l'Acadèmie Royale de Paris", Paris.

GARCÍA, A. (2002): "Legendre: La honestidad de un científico", Colección *"La Matemática en sus personajes"*, número 11, Nívola libros y ediciones, S.L., Madrid.

GAUSS, C.F. (1855): *Méthode des moindres carrés. Mémoires sur la combinaison des observations,* Traduits en Français et publiés avec l'autorisations de l'auteur, par J. Bertrand, Mallet-Bachelier, Paris.

HALD, A. (1990): *A History of Probability and Statistics and Their Applications before 1750,* Wiley, Nueva York.

HALLEY, E. (1693): "An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslau; with an attempt to ascertain the price of annuities upon lives", *Philosophical Transactions*, número 17, 596-610.

HALLEY, E. (1716): "Methodus singulares qua Solis parallaxis sive distantia a Terra, ope Veneris intra Solem conspiciendae tuto determinari poterit", *Philosophical Transactions*, número 348, 454-610.

HUYGENS, C. (1742): *De Ratiociniis in Ludo Alae*, Opera Varia, Lugduni Batavorum.

KRAMP, C. (1799): *Analyse des Réfractions astronomiques et terrestres,* Dannbach, Strasbourg and Schwikkeert, Leipzig.

LAPLACE, P.S. (1781): "Mémoire sur les l'inclinaison moyenne des orbites des cometes, sur la figure de la terre, et sur les functions", *Mém. Acad. R. Sci. Paris*, *(Savants Étrangers),* 7, 503-540.

LAPLACE, P.S. (1788): "Théorie de Júpiter et de Saturno", *Mém. Acad. R. Sci. Paris*, *1785*, 33-160.

LEGENDRE, A.M. (1806): *Nouvelles Méthodes pour la Détermination des Orbites des Comètes*, Chez Courcier, París.

MICHELL, J. (1767): "An inquiry into the probable paralax, and magnitude of the fixed stars, from the quantity of light which they afford us, and the particular circumstances of their situation", *Philos. Trans. R. Soc. London*, 57, 233-264.

MICHELL, J. (1784): "On the means of discovering the distance, magnitude, & c., of the fixed stars", *Philos. Trans. R. Soc. London*, 35-57.

RUIZ, G. (1990): "De cuando los matemáticos construían el universo", *Epsilon,* 16, 45-56.

RUIZ, G. (2003): "Los orígenes del método de mínimos cuadrados", Revista *SUMA*, número 43, 31-37.

SÁNCHEZ, C. Y VALDÉS, C. (2001): "Los Bernoulli: Geómetras y Viajeros", Colección *La Matemática en sus personajes*, número 10, Nívola libros y ediciones, S.L., Madrid.

Capítulo 14

The aftermath of Abraham de Moivre's Doctrine of Chances and Annuities on Lives in 18th-century Europe

IVO SCHNEIDER
Münchner Zertrum fúr Wissenschafts-und
Technikgeschichte
Deutches Museum

Some biographical data of De Moivre[i]

Abraham Moivre was born in May 26, 1667 as the son of a Protestant surgeon in Vitry-le-François in the Champagne. He spent the first 20 years of his life in France where he was educated in different Huguenot institutions until they were closed in the early 1680s. Quite early he had turned to mathematics. At 16 he had studied amongst other things Huygens' tract "De ratiociniis in ludo aleae". In Paris in the 1680s he was taught by the private teacher of mathematics Jacques Ozanam, who might have offered to Moivre a model for how to make his living when he had to support himself shortly afterwards.

After the revocation of the Edict of Nantes in 1685 hundreds of thousands of Huguenots left France. Amongst them was Moivre who went to England together with his brother Daniel in December 1687. In England he began his occupation as a tutor in mathematics. Here he added a "de" to his name. De Moivre mastered Newton's Principia from 1687 very early and became a true and loyal Newtonian. In 1692 he met Halley and Newton and in 1697 he was chosen a Fellow of the Royal Society. He was naturalized in 1705, and, when Newton had nominated him a member of the commission to decide the priority dispute between himself and Leibniz in favour of Newton in 1712, he gave up every hope of finding a position as a university professor for mathematics on the continent. But even in England he had to learn that for him the only way to make a living was to work as a private teacher and as a consultant for problems concerning different forms of annuities. In order to attract well-to-do clients he had to make a reputation as a mathematician. He soon found out that it was much easier to do that in the calculus of games of chance than in the new field of the infinitesimal calculus. Francis Robartes, one of de Moivre's clients, draw de Moivre's attention to the first edition of

Montmort's Essay d'analyse sur les jeux de hazard from 1708, which raised de Moivre's interest in the theory of games of chance and probability. In the Philosophical Transactions for 1711 de Moivre published a longer article De mensura sortis, seu, De Probabilitate Eventuum in Ludis a Casu Fortuito Pendentibus on the subject, which was followed by his Doctrine of Chances, the first edition of which was published in 1718. A second edition from 1738 contained de Moivre's normal approximation to the binomial distribution, which he had found in 1733. The third edition from 1756 contained as a second part the Annuities on Lives, which had been published as a monograph for the first time in 1725.

The Doctrine of Chances is in part the result of a competition between de Moivre on the one hand and Montmort together with Niklaus Bernoulli on the other. De Moivre claimed - very much to the annoyance of Montmort, but justified by his later work - that his representation of the solutions of the then current problems tended to be more general than those of Montmort. This holds at least for the second and third edition of the Doctrine of Chances, which offered so many new results that Montmort's contributions to the subject fell justly into oblivion.

Precursors of de Moivre

Christiaan Huygens with his tract De ratiociniis in ludo aleae[ii] of 1657 was the first author who acquainted de Moivre with the calculus of games of chance. In England de Moivre read a small booklet Of the Laws of Chance, or, a Method of Calculation of the Hazards of Game, published anonymously by John Arbuthnot in 1692, which was largely inspired by Huygens' tract. When in 1708 the Essai d'Analyse sur le Jeux de Hasard of the French nobleman Pierre Rémond de Montmort appeared anonymously, de Moivre hurried to read this work written in his native tongue.

Edmond Halley's "An Estimate of the Degrees of the Mortality of Mankind", published in the Philosophical Transactions for 1693[iii], might have become known to de Moivre quite early, since he had met Halley in 1692 in person. When de Moivre wrote De Mensura Sortis[iv], his first tract concerning games of chance and probability, he had no knowledge of Jakob Bernoulli's Ars Conjectandi, which was published in 1713, eight years after Bernoulli's death. In the preface of the Ars Conjectandi Nikolaus Bernoulli, Jakob Bernoulli's nephew, had asked de Montmort and de Moivre to continue his uncle's work, the application of the calculus of probabilities "to economical and political uses". Shortly after the Ars Conjectandi the second edition of Montmort's Essai of 1714 appeared, which contained, apart from some improvements, an extension of the first edition and the correspondence between Montmort and Nikolaus Bernoulli up to the end of 1713.

The long preface for the first edition of the Doctrine of Chances of 1718 shows that de Moivre understood the second edition of the Essai as a challenge to replace the solutions offered by Montmort and Nikolaus Bernoulli by his own. This is one reason why he showed no interest in continuing Jakob Bernoulli's project of applying the theory of probability to economics and politics as outlined in the Ars conjectandi.

Surprisingly, de Moivre seems to have been unfamiliar with the correspondence between Pascal and Fermat from 1654 concerning games of chance, part of which was published still in the 17^{th} century[v], and with Blaise Pascal's Traité du triangle arithmétique published in 1665. In constrast to Huygens, Pascal and Fermat used combinatorial methods for the solution of the problems discussed by them. But even Jakob Bernoulli, who had devoted the second part of the Ars Conjectandi to combinatorics, was obviously totally ignorant of Pascal's contributions to combinatorics.

De Moivre's main contributions to stochastics

De Moivre's main contributions to stochastics are contained first of all in his The Doctrine of Chances: Or, A Method for Calculating the Probability of Events in Play, with editions in 1718, 1738, and 1756 and in his Annuities on Lives, which appeared first in 1725; later editions came out in 1731, 1743, 1750, 1752 and together with the Doctrine of Chances in 1756.

The Miscellanea analytica of 1730 dealt inter alia with open issues in the controversy with Montmort who, however, had died already in 1719. In a supplement to the Miscellanea analytica de Moivre had started, in competition with James Stirling, to find an approximation for n-factorial when n is large. This led in 1733 to the first formulation of the central limit theorem in the Approximatio ad Summam Terminorum Binomii $(a+b)^n$ in Seriem expansi which de Moivre declared as a private communication for some friends as distinct from a publication. The publication of this result occurred only in the second and third edition of the Doctrine of chances[vi].

The main achievements contained in the Doctrine are a first theory of probability in the introduction[vii], which contains the principal concepts like probability, conditional probability, expectation, dependent and independent events, the multiplication rule, and the binomial distribution.

De Moivre had also developed algebraical and analytical tools for the theory of probability for instance a "new algebra" for the solution of the problem of coincidences, which forshadowed Boolean algebra[viii], the theory of recurrent series for the solution of difference equations, and the method of generating functions. He used the method of generating functions, for example, for solving the general dicing problem of finding the number of chances to get s points with a throw of n dice each having f faces[ix]. De Moivre determines the number by taking the coefficient of x^s in the development of

$$\left(\sum_{i=1}^{f} x^i\right)^n .$$

He had developed for the solution of the problem of the duration of play the theory of what he called recurrent series, which are in modern terminology homogeneous linear difference equations with constant coefficients.

De Moivre's greatest mathematical achievement is considered a form of the central limit theorem, which he found in 1733 at the age of 66. He understood his central limit theorem as a generalization and a sharpening of Bernoulli's main theorem in the fourth part of the Ars Conjectandi, which was later named the law of large numbers by Poisson[x].

Starting from Bernoulli's law of large numbers in the form

$$1 > P\left(\left|r_n - p\right| \le \varepsilon\right) > \frac{c}{c+1}$$

where r_n is the relative frequency of the occurrence of independent events E_i, $i = 1,\ldots,n$ with $P(E_i) = p$, ε arbitrarily small and c arbitralily large quantities in a sufficiently large number n of trials, de Moivre found on the basis of an approximation of log n! for large n and of Bernoulli's formula for the sum of powers of integers the equivalent of

$$\lim_{n\to\infty} P\left(|r_n - p| \le c\sqrt{\frac{2p(1-p)}{n}} \right) = \frac{2}{\sqrt{\pi}} \int_0^c e^{-t^2} dt.$$

De Moivre calculated the numerical values 0.682.., 0.954.. and 0.998..of this integral for c = 1, 2, and 3. He concluded from this result that "altho' Chance produces Irregularities" those irregularities shall cancel out with time in order to reveal the "Order which naturally results from Original Design" and so the existence of God, "the great Maker and Governour of all"[xi]. Thus he interpreted his central limit theorem as a proof for the existence of God, an interpretation, which constituted a challenge for theologians, especially if they were sceptical of natural religion.

Next to games of chance de Moivre considered problems dealing with human mortality as the proper subject of probability theory. He collected his results in this field in his Annuities on lives. His starting point for the Annuities on lives was Halley's table of the population of Breslau published in the Philosophical Transactions for 1693, which consisted of the numbers of those living at age 1 to over 80 and was interpreted as a mortality table. Halley's table induced de Moivre to assume a constant decrement of the living in every year after age 12, which means that the probability $p_x(m)$ to survive with age x the next m years is

$$\frac{l_{x+m}}{l_x} = \frac{a_{\max} - (x+m)}{a_{\max} - x} \quad \text{where} \quad 12 \le x < a_{\max},\ 0 \le m \le a_{\max} - x \quad \text{and} \quad a_{\max} = 86.$$

If i is the rate of interest and the amount paid out every year is 1, the present value of an annuity on lives for a person aged x is

$$a_x = \sum_{j=1}^{a_{\max}-x} (1+i)^{-j} \cdot p_x(j),$$

for which de Moivre found the following recurrent series:

$$a_x = (1+i)^{-j} \cdot p_x(1) \cdot (1 + a_{x+1}).$$

For the determination of the annuities of joint lives he used, for the sake of ease of calculation, the hypothesis of what he called fictitious lives. According to this hypothesis the sequence of the probabilities to stay alive for the next m years at age x is a geometrical series, that is to say:

$$\frac{l_{x+m}}{l_x} = \left(\frac{l_{x+1}}{l_x} \right)^m.$$

Because of the sometimes great differences between the annuities calculated with "fictitious" lives and annuities calculated on the basis of "real" lives this hypothesis was eventually given up.

The impact of de Moivre's main publications on relevant English publications

Since de Moivre's most important works in stochastics the Doctrine of Chances and the Annuities on lives were written in English his impact on British authors was by far the most intense and lasting. The English mathematician who perhaps owed most to de Moivre's work and took advantage of de Moivre's achievement already during de Moivre's lifetime was Thomas Simpson. Simpson worked like de Moivre as a teacher of mathematics; he wrote many mathematical textbooks amongst them two concerning probability theory and annuities. He plagiarized and further developed de Moivre's work in The Nature and Laws of Chance[xii] and in The Doctrine of Annuities and Reversions[xiii].

The nature and laws of chance of 1740 contains the results of the Doctrine of Chances - without the descriptions of the different games of chance -, the general solution of the problem of the duration of play and the determination of annuities on lives; although de Moivre is mentioned respectfully in the preface as the man who had developed the subject, his name does not appear in the text in order to mark the results taken from de Moivre's Doctrine. Simpson's book was much cheaper than de Moivre's, in order to attract buyers to the detriment of de Moivre.

Simpson's Doctrine of annuities and reversions follows to a great deal de Moivre's Annuities on lives but can claim for its second edition a better solution for annuities of joint lives compared with de Moivre.

Whereas Simpson's indebtedness towards de Moivre's work is obvious, the influence of de Moivre on Thomas Bayes' Essay, which was published posthumously in 1764[xiv], remains hypothetical, despite many possible direct or indirect relationships between de Moivre and Bayes, via Colin MacLaurin and the Earl of Stanhope[xv], or in meetings of the Royal Society of which both de Moivre and Bayes were fellows - relationships determined by mathematical and theological issues[xvi]. A possible motive for Bayes to write his Essay was Moivre's second remark with respect to his central limit theorem[xvii]:
"As upon the Supposition of a certain determinate Law according to which any Event is to happen, we demonstrate that the Ratio of Happenings will continually approach to that Law, as the Experiments or Observations are multiplied: so, conversely, if from numberless Observations we find the Ratio of the Events to converge to a determinate quantity, as to the Ratio of P to Q; then we conclude that this Ratio expresses the determinate Law according to which the Event is to happen."

Obviously de Moivre here addressed the inverse problem in the form to conclude from the outcome of many trials the unknown probability of an event. Since normally the possibility to undertake many trials is restricted or even non-existent Bayes became interested in the inverse problem when the number of trials may even be small:
"Given the number of times in which an unknown event has happened and failed: Required the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named."

De Moivre had also addressed all the topics dealt with in the published papers of Bayes. These are Bayes' first publication on God's benevolence[xviii], to which de Moivre's two remarks following his central limit theorem in 1738 and 1756 could be considered as a reaction. In addition there is a paper of Bayes published posthumously in the Philosophical Transactions for 1763[xix]. It deals with the (asymptotic) series for $\log n!$ in the form developed by de Moivre in 1730 in his Miscellaneis Analyticis Supplementum. De Moivre had touched upon the semiconvergent character of this series, but claimed in the end that it is convergent[xx]. The series was used by de Moivre in order to prove the central limit theorem in 1733.

Bayes' posthumously published papers were edited by Richard Price (1723-1791), a very prominent Presbyterian Minister, who became known at first by his publications on ethical and political questions. He was made a Fellow of the Royal Society in 1765 after his edition of the two papers of the late Thomas Bayes, which were published in the Philosophical Transactions in 1764 together with an introduction and some remarks and proofs by Price. In 1771 Price's Observations on reversionary payments with a long subtitle appeared. This book had several editions in a few years and remained a standard text for actuarial science for the next hundred years. In it references to de Moivre's contributions to annuities on lives appear throughout.

James Dodson (ca. 1705-1757), a former student of de Moivre, called himself in his major work the Mathematical Repository "Accomptant, and Teacher of the Mathematics". Dodson's Repository was published in three volumes in 1748, 1753, and 1755. The first volume was dedicated to De Moivre; the second contained as the first of three parts "Indetermined Questions, solved generally, by an elegant method communicated by Mr De Moivre". The two other parts of the second volume deal with questions "relating to Chances and Lotteries" as well as those "concerning Annuities for Lives". In the Preface to the second volume from 1753 Dodson defended de Moivre's "truly admirable hypothesis, that the decrements of life may be esteemed nearly equal, after a certain age"; which is to say "that the number of those who die within a year is constant or that the number of living after x, x+1, x+2, ... years constitutes an arithmetical series". When De Moivre's critics had mantained that his hypothesis does not hold for the recently available data for London for the years from 1744 to 1753, for which the Reverend William Brakenridge had calculated a new life table, Dodson found that de Moivre's hypothesis holds even in this case.

The Repository not only earned Dodson election to the Royal Society in January 1755, but also allowed him to make a reputation as an actuary who formulated the principles on which the Equitable Society, one of the first life insurance companies in England was founded in 1762 five years after his death. Richard Price, who survived Dodson, became a consultant for the Equitable Life Assurance Society. So through his two mediators, James Dodson and Richard Price, de Moivre's work on life insurance influenced the development of the first insurance society in England, which was based on mortality statistics and on mathematics.

Apart from authors like Simpson, Dodson, Price and Bayes, who can claim at least a modest place in the history of mathematics, there are many British authors in the 18th century, who, forgotten for a long time, refer to de Moivre and his results in probability theory in textbooks as well as in popular books on games of chance and annuities on lives. To these belongs at his time the very popular Edmund Hoyle - "according to Hoyle" - in his Essay towards making the Doctrine of Chances Easy to those who understand Vulgar Arithmetic only (1754) or in his Polite Gamester (1761). An example how the Doctrine was exploited is given by Edmund Hoyle (1671-1769). In Problem III of his Essay he asks for the number of trials necessary to make the probability that an event, the chances of its happening to its failing being as a : b, is at least 1/2.

$$\frac{b^x}{(a+b)^x} = \frac{1}{2} \Rightarrow x\left[\ln(a+b) - \ln b\right] = \ln 2\,,$$

$$\text{or for } q = \frac{b}{a} \quad \Rightarrow \quad x\ln\left(1+\frac{1}{q}\right) = \ln 2 \text{ and if } q \gg 1 \Rightarrow x \approx q\ln 2 \approx q\cdot 0.7 \qquad .$$

In his third query Hoyle asks "In how many Throws may you undertake upon an Equality of Chance, to throw two Sixes upon two Dice?" and gives the answer: "The whole Number of Chances upon two Dice are thirty-six, out of which there is but one Chance for throwing two Sixes; therefore, (according to Mr. Demoivre's Doctrine) you are to multiply 35 by 0.7, which solves the question."

The result 24.5 leads to the somehow irritating final statement of Hoyle: "You may undertake to do it in twenty-four Throws and a half."

In the same way Hoyle solved in chapter 8 the problem "In a Lottery to find out the Number of Tickets, which is requisite to entitle you to a Prize, upon an Equality of Chance" for different numbers of blanks and one prize.

Still in de Moivre's lifetime in 1741 Richard Hayes, an accomptant and writing master, had published his The gentleman's complete book-keeper. Its chapter 27 "Shews the Ways of casting up the Value of Annuities for successive Lives, and joint Lives, by Common Arithmetick, according to Mr. De Moivre's and Mr. Lea's Methods".

After his death most references to de Moivre in the English literature concern his work on life annuities. So James Ferguson (1710-1776), a lecturer on natural philosophy and inventor of mathematical instruments and FRS, published in 1767 his Tables and Tracts, Relative to Several Arts and Sciences. He refers to de Moivre's Annuities on Lives and the Doctrine for the solution of the problem[xxi]: "To find at what Rate of Interest I ought to lay out a sum S, so as it may increase1/3 for Instance, or become 4/3 S in 7 years." According to Ferguson "In questions concerning the Values of Lives any how combined, recourse must be had to Mr. De Moivre's last Edition of his Treatise on Annuities."

Another example for de Moivre's lasting authority concerning actuarial problems is offered by Robert Hamilton (1743-1829), a political economist and mathematician, who received in 1779 the chair of natural philosophy of Aberdeen University and in 1817 the chair of mathematics after he had taught mathematics at Aberdeen for many years. He published An Introduction to Merchandize, which came out in 1777 as the first of a whole series of practical treatises; in its first edition de Moivre is mentioned several times in chapter VIII of Part II which deals with "Annuities for lives" but not in the preceding chapter entitled "Doctrine of Chances". However, credit is given to De Moivre's general solution by a generating function of Hamiliton's problem V "Required the chance of throwing any proposed number with a given number of dice" in the chapter concerning the "Doctrine of Chances" in the second edition from 1788[xxii]. In the third edition from 1797 the same problem is solved by de Moivre's formula, however without mentioning his name. In the chapter on annuities there are again several references to de Moivre.

De Moivre is also mentioned by Hamilton in his Mathematical Tables from 1790 in "A Table of the value of an annuity of Life for a single life"[xxiii].

Francis Maseres (1731-1824), who came from a Huguenot family and whose plans to receive the Lucasian chair for mathematics in Cambridge failed, later worked as a lawyer and as a judge in London. He wrote a book The principles of the doctrine of life-annuities; explained in a familiar manner, so as to be intelligible to persons not acquainted with the doctrine of chances, which appeared in 1783. Maseres refered to de Moivre in this book very often, especially to de Moivre's hypothesis of the constant "decrements of human life". He was concerned with the differences between the annuities calculated according to de Moivre's hypothesis and those based on other assumptions.

As late as 1793 an anonymously published tract in Englisch on Faro and Rouge et Noir came out which contained an explanation of the two games together with a table showing the chances against the "punter" or Ponte in Faro or Pharao taken from de Moivre's Doctrine.

Most English 18th century publications use the expression "doctrine of chances" when refering to the theory of probability and so confirm even without mentioning de Moivre the influence on the subject of a man who had coined this expression. So Francis Maseres published in 1795 with The doctrine of permutations and combinations being an essential part of the doctrine of chances "as it is delivered by Mr. JAMES BERNOULLI, in his excellent Treatise on the Doctrine of Chances, intitled, Ars Conjectandi, and by the celebrated Dr. JOHN WALLIS, of Oxford, in a Tract intitled from the Subject, and published at the end of his Treatise on Algebra."

During de Moivre's lifetime nobody would have failed to name de Moivre when talking about the doctrine of chances, as one can see for instance from letters of the Abbé Jean Bernard Le Blanc which were originally published in French and appeared in English as Letters on the English and French nations, containing curious and useful observations on their constitutions natural and political in two volumes in 1747. Le Blanc reports[xxiv]: "I have had several conversations upon this subject with the famous M. DE MOIVRE, the greatest calculator of chances now in England: but I did not perceive that he had ever calculated the effects of gaming, with regard to morality, though that is a much more essential thing than the theory of chances."

He continues a bit later[xxv]: "I must add that the great gamesters of this country, who are not usually great geometricians, have a custom of consulting those who are reputable calculators upon the games of hazard. M. DE MOIVRE gives opinions of this sort every day at Slaughter's coffee-house, as some physicians give their advice upon diseases at several other coffee-houses about London."

One should add that le Blanc is an involuntary witness of de Moivre's prominence in the calculus of probabilities in England, since he was much more concerned with the bad effects of gaming, which makes according to him every gamester a loser.

De Moivre's influence on the continent

De Moivre influenced the development of mathematics outside Great Britain much less than in his second homeland for different reasons. For the Bernoullis, who had taken sides with Leibniz in the priority dispute with Newton, de Moivre stood on the wrong side. And even if de Moivre was considered by some as the leading mathematician after Newton in England, his mathematical output seemed modest compared with that of the leading mathematicians in continental Europe in the generations after Leibniz. Most of the continental mathematicians, who published mainly in Latin or French, were not prepared to read English, the language in which de Moivre's main works were published. None the less there are direct references to de Moivre's work in probability calculus and traces of its impact in continental Europe.

The most prolific mathematician in the 18th century, Leonhard Euler, certainly knew at least part of de Moivre's work, but we find nearly no direct hints to de Moivre in Euler's comparatively small oeuvre concerning probability theory[xxvi]. Euler was certainly an honest man but he was not very good in referring to the work of others. Euler's contributions to stochastics were to a great extent stimulated by requests from outside; as soon as he was confronted with a problem he would try to solve it no matter if others had solved it already. An example for this is the problem of finding the probabilities of two players engaged in the game of Rencontre which Maupertuis had posed to Euler. Euler hurried to answer Maupertuis the next day beginning with the remark[xxvii]: "I found the solution so difficult that I doubt very much that a more difficult problem was ever solved". Obviously Euler did not know, or did not care to know, that solutions were found years before by Montmort, Niklaus Bernoulli, and de Moivre. There are no direct hints to the Doctrine or the Annuities in Euler's works. He was probably familiar with the Mensura Sortis and certainly with the Miscellanea Analytica; this one can assume from Euler's terminology and references to the Miscellanea Analytica in his work. Similarity (and difference) in the treatment of several problems concerning lotteries, the coincidence (rencontre) in the game of Treize, the chances in Pharao, the duration of play, annuities on lives and human mortality and many others in his notebooks, like the determination of the probability of getting s points by throwing m dice each with n sides can, but are not necessarily to be interpreted as some familiarity with the relevant works of de Moivre.

De Moivre's impact on the further development of the analytical tools used in probability theory by Lagrange and Laplace is more easily proved, because both refer to de Moivre. In addition, Lagrange and Laplace tried, presumably independently of one another, to translate the Doctrine into French[xxviii], a project that apparently was never finished. Laplace refers in the introduction to the Théorie analytique des probabilités to de Moivre's determination of the general term of a recurring series and its summation as the first step in a development leading to Lagrange's recurro-recurrent series and his own theory of generating functions[xxix]. In his "note historique" Laplace honours even more de Moivre's extension of Bernoulli's law of large numbers and his contributions to annuities on lives.[xxx]

Whereas it is easy to disappear in the shadow of mathematical celebrities like Euler, Lagrange, or Laplace, because they are creative enough not to depend on the ideas of others, the less gifted are perhaps more prone to use and to refer to – if they are honest – the work of predecessors. For this group of minor figures I shall give two examples: The first concerns the two Cistercian Fathers Don Roberto Gaeta and Don Gregorio Fontana who translated de Moivre's Annuities on Lives into Italian and published it in 1776 in Milano as La Dottrina degli Azzardi Applicata ai Problemi della Probabilita della Vita, delle Pensioni Vitalizie, Reversioni, Tontine, ec. Di Abramo Moivre. This very well-informed book was used for a course in mathematics at the university of Pavia.

The book contains the most recent tables concerning human mortality and the numerous notes refer to the literature published after de Moivre; it contains the best bibliography available at the time, which includes the contributions of Euler, Süßmilch and Lambert to the subject. In addition one finds allusions to the Doctrine of Chances concerning the vignette on its first page and de Moivre's attempt to prove the existence of God.

The second example pertains to German probabilists stemming from the philosophical school of Christian Wolff. Representative for this group is Ludwig Martin Kahle, who in his Elementa Logicae probabilium methodo mathematica in usum scientiarum et vitae adornata[xxxi] refers in the preface to the reader to de Moivre's Mensura sortis.

---

[i] For more extensive biographical information see Ivo SCHNEIDER, Abraham de Moivre (1667-1754). In: Oxford Dictionary of National Biography (hrsg. v. H.C.G. MATTHEW und Brian HARRISON) vol. 38, Oxford 2004, S. 522 f. and Ivo SCHNEIDER, Der Mathematiker Abraham de Moivre (1667 bis 1754). In: Archive for History of Exact Sciences 5, 1968, S. 177-317.

[ii] Christiaan HUYGENS, De ratiociniis in ludo aleae, in: Frans VAN SCHOOTEN, Exercitationum mathematicarum libri quinque, Leiden 1657, p. 517-534.

[iii] Edmund HALLEY, An estimate of the Degrees of the Mortality of Mankind, drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives, in: Philosophical Transactions Nr. 196 (1693), p. 596-610.
[iv] Abraham DE MOIVRE, De mensura sortis, seu, de probabilitate eventuum in ludis a casu fortuito pendentibus, in: Philosophical Transactions Nr. 329 (1711), p. 213-264.
[v] Pierre DE FERMAT, Varia opera mathematica, Toulouse 1679.
[vi] Doctrine of chances, London 1738, p. 235-243, and London 1756, p.243-254.
[vii] Doctrine 1738, p. 1-30, 1756, p. 1-33.
[viii] Doctrine 1718, p. 61-63, 1738, p. 96-98, 1756, p. 110-112.
[ix] Doctrine 1738, p. 37-39, 1756, p. 41-43.
[x] S.-D. POISSON, Recherches sur la probabilité des jugements en matière civile, Paris 1837, p. 7.
[xi] Doctrine 1738, p.243, 1756, p. 251 and 252.
[xii] London 1740.
[xiii] London 1742, second edition London 1775.
[xiv] Thomas BAYES, An essay towards solving a problem in the doctrine of chances, in: Philosophical Transactions vol. 53 (1763), p. 370-418.
[xv] See David BELLHOUSE, The Reverend Thomas Bayes, FRS: A Biography to Celebrate the Tercentenary of His Birth, in: Statistical Science 19 (2004), no. 1, p. 3–43.
[xvi] See Ivo SCHNEIDER, De Moivre's limit theorem and its possible connection with Bayes' Essay. In: Physica et historia – Festschrift für Andreas Kleinert zum 65. Geburtstag (hrsg. von Susan SPLINTER, Sybille GERSTENGARBE, Horst REMANE und Benno PARTHIER) (= Acta Historica Leopoldina, Nr. 45), Halle (Saale) 2005, p. 155-161.
[xvii] Doctrine 1756, p. 251.
[xviii] Thomas BAYES, Divine Benevolence, or an attempt to prove that the principal end of the divine providence and government is the happiness of his creatures. London 1731.
[xix] A letter from the late Reverend Mr. Thomas Bayes, F.R.S. to John Canton, M.A. and F.R.S. in: Philosophical Transactions vol. 53 (1763), p. 269-271.
[xx] Abraham DE MOIVRE, Miscellaneis analyticis supplementum, London 1730, p. 9.
[xxi] p. 286.
[xxii] p. 183.
[xxiii] p. 126.
[xxiv] vol. II, p. 307.
[xxv] vol. II, p. 309.
[xxvi] Ivo SCHNEIDER, I contributi di Euler alla stocastica nel contesto della letteratura contemporanea. In: Quaderni della Accademia delle Scienze di Torino, 2008, p. 103-121.
[xxvii] Letter from Euler to Maupertuis of April 16, 1752. See: Leonhard Euler Opera omnia IV, 6, Basel 1986, p. 203 f.
[xxviii] See letter from Lagrange to Laplace of December 30, 1776. In: Oeuvres de Lagrange vol. 14, Paris 1892, p. 66.
[xxix] See Pierre Simon LAPLACE, Théorie analytique des probabilités, third edition, Paris 1820, p. XXIX.
[xxx] Ibidem p. CXLVII.
[xxxi] Published in Magdeburg 1735.

Capítulo 15

Justificaciones de la regla de Borda: Una revisión crítica

MIGUEL MARTÍNEZ PANERO
Universidad de Valladolid

Introducción

En 1770 el *chevalier* Jean-Charles de Borda, un reputado ingeniero militar y oficial de marina con aportaciones a la dinámica de fluidos y al perfeccionamiento de instrumentos náuticos y cartográficos[1], expuso ante la Academia de Ciencias de París una memoria ajena a sus intereses científicos primordiales, pero que, en buena medida, le ha reportado su fama póstuma. En calidad de miembro de la citada institución, Borda había constatado que el método de pluralidad empleado para elegir a los nuevos miembros entre los aspirantes a las distintas secciones (Mecánica, Astronomía, etc.) podía tener efectos perversos[2]. Por ello, redactó una propuesta justificada de enmienda de tal procedimiento de selección, propugnando en su lugar el que hoy lleva su nombre, que sería implantado en la Academia[3] desde 1796 hasta 1804, fecha en que dejó de ser utilizado por indicación expresa de Napoleón[4].

---

[1] Especial atención merece su círculo repetidor, que tuvo un papel primordial en la implantación del sistema métrico decimal.
[2] Sobre los defectos en que incurre la regla de pluralidad, véanse la propia memoria de Borda [1770](1784) o Morales (1797). Un análisis crítico más reciente (pues es aún usada hoy en día en numerosos contextos decisionales) aparece en García Lapresta – Martínez Panero (2000).
[3] Denominada "Instituto Nacional de Francia" tras el periodo revolucionario.
[4] Curiosamente, en 1997 Napoleón había sido elegido miembro de la sección de Ciencias Físicas y Matemáticas, subsección de Artes Mecánicas, mediante el método de Borda.

Condorcet parafraseó como sigue el método de Borda, para criticarlo a su vez (citamos por la traducción[5] de López de Peñalver [1799] (1992, p. 57)):

Lo prolijo de la cita (muy en el estilo de la época) queda compensado por su innegable interés histórico y científico. Efectivamente, en la citada exposición se consideran distintos juegos de ponderaciones o pesos para implementar la regla de Borda (si bien equivalentes, en el sentido de que todos determinan la misma decisión colectiva), cuyo espectro, desde su misma propuesta hasta nuestros días, ha sido cuestionado por su presunta arbitrariedad[6].

En el presente trabajo se hace un repaso crítico de las justificaciones de tal rango de puntuaciones empleadas en la regla de Borda. Así, en la Sección 2, se enuncian los razonamientos del propio Borda o de Laplace, quienes preestablecen equidistribuciones de valores a asignar con criterios probabilísticos o estadísticos. Así mismo, se expone un argumento alternativo de Borda (secundado independientemente por Morales) que basa las puntuaciones otorgadas en una forma compacta de tener en cuenta simultáneamente todas las victorias por mayoría simple en las comparaciones por parejas de las alternativas. A continuación, en la Sección 3, se exponen distintas justificaciones axiomáticas del método, en las que diversas combinaciones de propiedades deseables desde supuestos ético-democráticos determinan que la única regla de votación[7] que las satisface es justamente la de Borda. En la Sección 4 se trata su redescubrimiento por Kendall y desarrollos ulteriores basados en tratamientos métricos y de cercanía al consenso, así como con metodología DEA. La panorámica presentada culmina en la Sección 5, donde se esboza el razonamiento de tipo geométrico con el que Saari se ha manifestado recurrente e incontrovertiblemente a favor de dicha regla. Por fin, se concluye el trabajo con algunas consideraciones y posibles extensiones del marco de preferencias utilizado.

Borda, Laplace, Morales: apologías[8] del método

Aunque se suelen datar los orígenes de la regla de Borda a finales del siglo XVIII, estos deberían retrotraerse varios siglos más si tenemos en cuenta que Nicolás de Cusa, en el libro III de su obra *De Concordantia Catolica*, publicada en 1434, ya discute en un contexto electoral un método que es justamente el de Borda, con puntuaciones individuales escalonadas, desde la unidad hasta el número total de candidatos, asignadas a los mismos de

---

[5] Entre 1798 y 1799, se editaron en castellano las *Cartas a una Princesa de Alemania sobre diversos Temas de Física y Filosofía*, obra de divulgación de Leonhard Euler, en traducción de López de Peñalver. Esta versión incorpora la apostilla, que Condorcet incluyó en su edición de la obra de Euler, en la que el autor francés reseña la Memoria que Borda presentó ante la Real Academia de Ciencias de París, donde éste defendía su método de votación con puntuaciones.

[6] Es inmediato verificar que, para una misma situación preferencial, el ganador puede depender del juego de ponderaciones utilizado. En Stein – Mizzi – Pfaffenberger (1994) se analiza mediante una condición de dominancia la posibilidad de que una alternativa sea ganadora cualquiera que sea el sistema de puntuaciones utilizado.

[7] A lo largo del texto se emplearán recurrentemente los conceptos de *regla de votación*, *función de bienestar social* y *función de elección social*, para cuyo significado y diferencias según autores remitimos a Taylor (2005, pp. 7–8).

[8] Empleamos esta palabra en su significado de defensa o (auto)justificación, que procede de su etimología griega y se ha conservado más en el ámbito anglosajón que en el castellano (así, por ejemplo, en la célebre *Apología de Sócrates* de Platón o en *A Mathematician´s Apology* de Hardy).

forma secreta[9]. Ésta es la razón por la que McLean – Urken (1995, pp. 24) han reivindicado su prioridad y sugerido que debería hablarse de “regla de Cusa”, en vez de “regla de Borda”.

Sin embargo Cusa se preocupó más de aspectos meramente procedimentales que del trasfondo del método, al contrario que Borda. Y es que, si bien la memoria de este último, un texto breve y programático, tiene fines eminentemente pragmáticos (implementar un método idóneo para ser usado en el seno de la Academia de Ciencias de París), su incisión y aguda visión del problema electoral le ha hecho merecer un destacado puesto en la historia de la Teoría de la Elección Social.

Borda [1770] (1784, p. 659), como ya se ha indicado, postula una cardinalización equidistante del ordenamiento de alternativas utilizando lo que Black (1958, p. 182) ha denominado “teoría de probabilidad a nivel de sentido común”:

> *“Digo que el grado de superioridad que este elector concede a A sobre B debe considerarse el mismo que el que otorga a B sobre C. En verdad, ya que el segundo candidato B es igualmente susceptible de todos los grados de mérito entre los de los otros candidatos A y C, no hay razón para sostener que el elector que ha decidido este orden de los 3 candidatos haya deseado colocarle más cerca de A que de C.”*

Las cursivas son nuestras, e indican una de las críticas de Condorcet al método de Borda (véase López de Peñalver [1799] (1992), p. 59):

> *“[Borda] supone que los grados de mérito o de preferencia entre los candidatos son iguales y reputados por iguales en la opinión de los censores. Esto admitiría bastante contradicción, y las más de las veces no puede suponerse esta igualdad, a lo menos en la opinión de los electores”.*

También Grazia (1953) hace recaer una de las debilidades del razonamiento de Borda en la cita anterior, toda vez que el argumento negativo de ausencia de información, aunque puede tener sentido en el marco de la memoria,

> *“[…] no le faculta para justificar lo contrario [esto es, que los pesos asignados deben ser exactamente proporcionales a las posiciones]. Y esto es exactamente lo que Borda supone hasta el punto de distorsionar la escala electoral.”*

Laplace dió una justificación estadística más sofisticada (aunque también controvertida, como veremos) al requerimiento de Borda[10]. Es bien conocido el carácter elíptico de los razonamientos de Laplace. Tanto es así que, en ocasiones, sus palabras: “Es fácil comprobar…” (o similares) requieren arduos cálculos por parte del lector para corroborar la veracidad de lo afirmado. Esto es exactamente lo que ocurre en este caso, en el que Laplace [1814] (1987, pp. 114–115) sostiene:

> *“Supongamos que a cada elector se le da una urna conteniendo una infinidad de bolas, por medio de las cuales poder matizar todos los grados de mérito de los candidatos;*

[9] El capítulo en el que Cusa expone su propuesta de la regla de Borda aparece traducido al inglés en McLean – Urken (1995, pp. 77–78).

[10] Aunque la cita de Laplace que sigue está tomada de su *Ensayo Filosófico sobre las Probabilidades*, publicado en 1814, según Black (1958, pp. 181–182) los argumentos expuestos ya aparecían en 1795 en las lecciones impartidas por el autor francés en la Escuela Normal, tal como indica el propio Laplace en el prólogo de dicha obra. Por otra parte, Besio y Banfi indican en las notas a su edición del *Ensayo Filosófico* que éste se publicó simultáneamente como introducción de la segunda edición de la *Teoría Analítica de las Probabilidades*, si bien su redacción definitiva corresponde a la quinta edición de 1825, que ha sido la referencia utilizada para realizar su traducción.

supongamos además que saca de la urna un número proporcional a los méritos de cada candidato […]. Pero las papeletas [en las que sólo se ordenan a los aspirantes] lo único que indican es que el primero tiene más [bolas] que el segundo, el segundo que el tercero, y así sucesivamente. Suponiendo que […] todas las combinaciones […] son igualmente admisibles, […] si estos números son muy elevados, como es de suponer que ocurra para que puedan expresar toda la gama de méritos, el análisis más elemental permite ver que los números que han de ser escritos en cada papeleta al lado del último nombre, del penúltimo, etc., pueden ser representados por la progresión aritmética 0, 1, 2, etc[11]*.”*

Todhunter (1865, pp. 546–548) reconstruyó el "elemental" razonamiento sugerido por Laplace para obtener la discretización indicada como sigue:

"Sean $t_1, t_2, \ldots, t_n$ *los méritos de los candidatos, empezando por el más meritorio […]; el elector puede adscribir cualesquiera méritos a los aspirantes, ordenados de tal forma que ninguno sea mayor que el inmediatamente precedente, y a lo sumo con un valor a.*

El valor medio del mérito del r-ésimo candidato será:

$$\frac{\iiint\cdots t_r dt_1 dt_2 \ldots dt_n}{\iiint\cdots dt_1 dt_2 \ldots dt_n}.$$

El registro de condiciones de integración efectuado por Laplace no es inteligible, y establece el resultado de integración sin explicación."

Hasta aquí, todo está esencialmente en Laplace, quien en el Libro 2, pp. 277–278, de su *Teoría Analítica de las Probabilidades*[12], llega sin justificación al valor $\dfrac{(n-r+1)a}{n+1}$,

el mismo que obtiene Todhunter apoyándose en los teoremas de cambio de variable y de Lejeune-Dirichlet. De aquí se sigue que cada elector debería asignar el número *n* al candidato considerado el mejor, el número *n*-1 al siguiente, y así sucesivamente, resultando ganador el aspirante que obtenga mayor suma de puntuaciones. Es decir, se reproduce exactamente el método propugnado por Borda.

Cabe señalar que Laplace es consciente del sesgo que tomaría la elección idónea cuando los agentes rebajasen estratégicamente el mérito de unos candidatos en favor de otros[13].

Y también señala otra crítica recurrente en la literatura que le ha sido hecha al método de Borda: su susceptibilidad de seleccionar candidatos mediocres[14].

---

[11] En realidad Laplace escribe "1, 2, 3, etc." (y así lo recogen Besio y Banfi en su edición). Sin embargo el matiz de traducción no incurre en error, pues ambas graduaciones, al diferir en una constante, proporcionan esencialmente el mismo método. De hecho, algo parecido ocurrió con Morales (1797), quien empleó la escala 1,2,3,… para cambiarla en Morales (1805) por la secuencia 0,1,2,… por las razones que se concretarán más adelante.

[12] Tomo VII de la edición de las Obras Completas publicado en 1886. Puede consultarse online en edición facsímil en http://math-doc.ujf-grenoble.fr/cgi-bin/oeitem?id=OE_LAPLACE__7_R2_0.

[13] También Condorcet (1785), en su *Discours Peliminaire,* sin citar a Borda (hace alusión veladamente a un "geómetra célebre") pero refiriéndose sin género de duda a su método, reputa a éste de cabalístico. Tal vez se entienda mejor este comentario si se tiene en cuenta que en francés *cabale* tiene una acepción de "intriga o maniobra oculta", de la que carece el término "cábala" en castellano. Cuando este punto débil le fue señalado a Borda, éste contestó lacónicamente que su procedimiento sólo era válido para ser utilizado por hombres honorables (Black (1958, p.182).

Daunou (1803) fue quizá el primero en cuestionar el razonamiento de Laplace, basado en intensidades de preferencia, sobre método de Borda, pues además de los aspectos estratégicos señalados, según su lectura se obtendría una escala de puntuaciones en progresión geométrica y no aritmética. Esta es interpretación la de Daunou[15]:

> *"¿Qué representan los números de Borda? Simplemente, según el ciudadano Laplace, que el primer candidato obtiene más bolas que el segundo, el segundo que el tercero, y así sucesivamente. […] Laplace señala que cualquiera que sea el número de bolas que obtenga el primer candidato en una papeleta dada, todas las combinaciones de números menores que satisfagan las condiciones precedentes [i.e., los votantes actúan de buena fe y asignan las bolas de forma precisa] son igualmente admisibles. Y él añade que 'el número de bolas obtenido por cada candidato se puede encontrar realizando el total de todos los números asignados a él y dividiendo ese total por el número total de combinaciones'. O sea, que a cada candidato se le asignará el término medio de los valores asignados respecto del candidato inmediatamente superior. De manera que si el primer candidato obtiene [n=] 20, el segundo tendrá 10. En verdad, ya que a este segundo candidato se le asignaría alguno de los términos de la serie del 1 al 19, que suman 190, y al dividirse este valor por el número de combinaciones, que son 19, obtenemos 10 como cociente. […] De manera similar serían n/4 para el tercer candidato, n/8 para el cuarto, y así sucesivamente".*

Para Black (1958, pp. 182–183) la justificación estadística de Borda – Laplace es "insostenible" porque, al interpretar las puntuaciones otorgadas con un sentido utilitarista, se llega a que "un votante considere los méritos relativos de dos candidatos relacionados en alguna razón de tipo 5:4 ó 6:1, o alguna otra. Pero podemos sentirnos bastante seguros de que la mente humana no opera de esta forma". Además, señala Black, se incurre en el problema de la "comparación interpersonal de utilidades[16]". Por su parte, Saari (1995, p. 20) tacha tal explicación de "filosófica" e "insatisfactoria". De hecho, son varios los autores que coinciden en que el enfoque probabilístico dado a la Teoría de la Elección Social por Condorcet y Laplace en buena parte ha lastrado el desarrollo de la disciplina en Francia hasta su renacimiento, ya avanzado el siglo XIX, en el ámbito anglosajón.

No obstante, un tratamiento estadístico *a la Laplace*, que llena las lagunas explicativas de éste con herramientas formales modernas, ha sido retomado en Tanguiane (1991, pp. 80 y ss.) y Tangian (2000), quien obtiene las puntuaciones Borda normalizadas como los valores medios esperados (o esperanza matemática) de un vector *n*-dimensional aleatorio equidistribuido en el intervalo unidad[17]. Así mismo, Tangian (2000) sostiene que las puntuaciones enteras del método de Borda representan las utilidades cardinales de los agentes (sus "estimaciones latentes") de manera suficientemente adecuada, de modo que para electorados grandes el modelo de Laplace (preciso, pero inmanejable) se aproxima al de

[14] Aunque el defecto puede devenir virtud, denominada en este caso "respeto por la media". Véase a este respecto Black (1958, p. 56), quien afirma que "el candidato que debe ser elegido es el mejor situado, por término medio, en las relaciones de preferencias de los votantes".

[15] Véase McLean – Urken (1993, pp. 262 y ss), de donde traducimos.

[16] Sobre este controvertido problema, que cuestiona la validez misma de la Teoría de la Elección Social, véase Martínez Panero (2004, pp. 142–150) y las referencias allí indicadas.

[17] Tangian (1991) se basa en el análisis orden-estadístico de Kendall – Stuart (1969, pp. 268 y ss.) y observa, muy atinadamente, que las puntuaciones Borda normalizadas obtenidas son justamente las coordenadas del vector de centros de gravedad de las utilidades cardinales de los agentes, algo que está implícito en los argumentos de Laplace.

Borda (grosero, pero manejable)[18]. Por otro lado, la regla de Borda fue "redescubierta" en un contexto estadístico por Kendall (1962), quien trató la agregación de rankings como un problema de estimación y propuso una solución de consenso que maximizaba el acuerdo entre los votantes[19]. Además, una conexión entre la regla de Borda y una "regla de la media" obtenida desde supuestos probabilísticos puede encontrarse en Intriligator (1973). Y recientemente Heckelman (2003) ha diseñado una "regla de Borda probabilística" que asigna los pesos mediante un mecanismo de lotería.

Cabe indicar que tal vez Borda [1770] (1784) previese críticas ulteriores a la forma de equidistanciar las puntuaciones de su método debido a lo poco concluyente de su razonamiento anterior, pues ya en su misma memoria proporcionaba una segunda vía justificativa para introducir los pesos utilizados. A continuación exponemos los argumentos con los que Borda defendió esta segunda forma de entender su regla siguiendo la exposición del matemático ilustrado José Isidoro Morales, quien, al parecer, los redescubrió independientemente.

Es interesante la forma en que Morales tuvo conocimiento del método de Borda, que suscitó por parte del autor español una memoria (Morales (1797)) en la que también éste defendía el nuevo procedimiento frente a otros más al uso como, por ejemplo, el de pluralidad (véase Martínez Panero – García Lapresta (2003)). A través del periódico francés *La Décade Philosophique* Morales había sabido por una breve reseña que los miembros del Instituto Nacional de Francia habían escogido cinco plazas vacantes utilizando un método de elección que él llama "de compensación y suma" (no menciona, por no conocerlo entonces, el nombre de Borda). Tal procedimiento le parece idóneo, justo y exacto para reflejar la opinión que los votantes, y en su memoria, también con fines prácticos de implementación efectiva, Morales lo describe así: cada elector ordena de mejor a peor los c candidatos. De cada ordenación se asignan c puntos al primer candidato, c-1 puntos al segundo candidato,..., 2 puntos al penúltimo, y 1 punto al último candidato. Se suman los puntos recibidos por cada candidato y se escoge al que ha recibido más puntos.

El autor español divulgó su obra en distintos ámbitos nacionales y extranjeros, de manera que llegó a ser conocida y valorada por el mismo Borda. No obstante, ocho años después, nuestro autor escribe un apéndice (Morales (1805)) a la obra anterior con el fin de contestar a reparos y objeciones recibidas. Morales detalla que se ha criticado su método de compensación y suma por la rigidez en la asignación de c, c-1,...,2, 1 puntos, ya que ésta no permite reflejar con exactitud la intensidad en la opinión, es decir, constriñe la libertad del votante y no refleja el mérito del votado. Argumenta entonces, para fundamentar estas puntuaciones, que en las elecciones binarias (con sólo dos candidatos) la mayoría simple es el método adecuado para elegir y en el hecho de que, cuando hay más de dos candidatos, se podrían hacer elecciones binarias entre cada uno de los posibles pares (en total, $c(c$-1)/2 elecciones binarias) y luego sumar los votos que cada candidato obtuvo. Ahora bien, tal información numérica queda de hecho capturada en las puntuaciones obtenidas por el método de Borda: basta con que se emplee una escala c-1, c-2,...,2, 1, 0, equivalente a la utilizada en Morales (1797). En resumen, Morales (1805, p. 23) deja así el estado de la cuestión:

---

[18] Para Tangian (2000) "la anterior suposición [i.e., la presunción de Laplace de que las estimaciones de utilidad ordenadas de los agentes están equidistribuidas, una hipótesis tradicional en probabilidad y estadística en ausencia de información] implica en particular que la distribución de utilidades cardinales es independiente de las preferencias ordinales".

[19] Como veremos, el enfoque de Kendall está en la base de algunos de los tratamientos métricos de la regla de Borda que se expondrán en la siguiente sección.

"[…] Si se hace una elección por el orden de mérito, con la progresión de los números naturales, empezada desde cero, y de tantos términos cuantos sean los candidatos, hay una absoluta identidad entre este método de votar ó de elegir [la regla de Borda] y el de la elección binaria repetida tantas veces como combinaciones admite el número de candidatos […] También será cierto que este método gozará de la misma exactitud y rigor que la elección binaria, que se toma por norma de toda de toda elección justa. Porque el tal método no es otra cosa que una elección binaria repetida, y el mérito de su invención es haber ahorrado esa repetición".

Massó (2003) estima que ésta es

"una maravillosa justificación (ordinal) de su método basado en las puntuaciones […]. La argumentación de Morales parece un antecedente a la justificación del axioma moderno de consistencia usado en las caracterizaciones axiomáticas de la mayoría de los conceptos de solución propuestos por la teoría de los juegos cooperativos (núcleo, valor de Shapley, nucleolo, etc.)".

Y como también veremos, una noción de consistencia, si bien en el contexto de la Teoría de la Elección Social, ha resultado ser crucial en la más conocida de las caracterizaciones del método de Borda que se detallan a continuación.

Goodman – Markowitz, Gärdenfors, Young, Black: jusificaciones axiomáticas

Arrow [1963] (1974, pp. 202–203), en un apartado histórico a modo de apéndice, se refiere de esta forma a los problemas que plantea la regla de Borda, ya señalados por Black (1958):

"[…] este método confiere ponderaciones iguales a las diferencias de rango de los candidatos y, así mismo, da la misma ponderación a los distintos votantes. Lo primero plantea el problema de la mensurabilidad de la utilidad; lo segundo, el problema de las comparaciones interpersonales. Borda justifica el primer paso con un argumento esencialmente basado en la ignorancia, [y el segundo] sobre la base de la igualdad de los votantes. Estos temas han continuado repitiéndose. El razonamiento de Goodman – Markowitz (1952) puede considerarse, en efecto, como la [primera] justificación axiomática de la posición de Borda".

El objetivo de Goodman – Markowitz (1952) fue modificar las condiciones del célebre teorema de imposibilidad de Arrow para enunciar resultados positivos. Así, demuestran que la regla de Borda es la única función de bienestar social que, además de los usuales y comúnmente admitidos principios de optimalidad fuerte de Pareto y simetría (anonimato), cumple una tercera condición que requiere que "la implicación [en el agregado] que supone el cambio de un nivel [en la consideración de un candidato] al siguiente sea la misma cualquiera que sea el nivel de partida".

El carácter *ad hoc* de esta última condición (que de hecho se puede entender como una relectura del requerimiento de equidistancia de Borda) en cierta forma ha relegado este resultado frente a otras axiomatizaciones del método recurrentemente citadas en la literatura. Por ejemplo, la proporcionada por Gärdenfors (1973), quien establece que la regla de Borda es la única función de votación representable[20] que satisface neutralidad (tratamiento

[20] Una función de votación representable es básicamente una regla de puntuación (*scoring rule*), que se define mediante un vector de pesos con tantas componentes como alternativas de manera que cada agente asigna la puntuación s_1 a la alternativa que ocupe el primer puesto, s_2 a la segunda, y así sucesivamente

simétrico de las alternativas), optimalidad débil de Pareto y una tercera condición denominada independencia posicional fuerte. Esta última propiedad se puede enunciar así: siempre que una alternativa no ocupe alguno de los niveles intermedios entre dos alternativas *x* e *y* dadas, y tampoco empate con ellas, la relación social entre *x* e *y* no se verá afectada por ella.

El trabajo de Gärdenfors citado entiende la regla de Borda como un procedimiento esencialmente ordinal[21], y la defiende de la "acusación injusta de ser un método (estúpido) de amalgamar intensidades de preferencias [de los agentes mediante números impuestos]". Su importancia principal está en delimitar dos visiones irreconciliables del voto[22]: la de Borda (posicionalista) y la de Condorcet (no posicionalista). Sin embargo, tampoco su caracterización de la regla de Borda es la más reconocida. Tal mérito le corresponde a Young (1974), quien probó que el método citado corresponde a la única función de elección social que satisface neutralidad, fidelidad, cancelación y consistencia. Young define la propiedad de fidelidad como el hecho de que "socialmente más preferido" e "individualmente más preferido" signifiquen lo mismo cuando la sociedad es unipersonal. Así mismo, la propiedad de cancelación establece que, si para cualquier par de alternativas, la preferencia de una frente a la otra por parte de un agente se puede compensar con la preferencia inversa por parte de otro agente, entonces todas las alternativas son ganadoras. Por fin, el principio de consistencia o refuerzo (*reinforcement*) se puede describir como el requerimiento de que si dos coaliciones por separado dan lugar a la misma decisión social, entonces cuando se unen en un solo bloque tal decisión se preserve.

En un trabajo posterior, Young (1975) caracterizó de nuevo la regla de Borda como la única de las reglas de puntuación (que, a su vez, vienen dadas por funciones de elección social que cumplen anonimato, neutralidad, consistencia y continuidad[23]) que además cumple la mencionada propiedad de cancelación. Desde entonces se han proporcionado numerosas axiomatizaciones (en ocasiones más accesibles): Fine – Fine (1974), Hansson – Sahlquist (1976), Fishburn – Geherlein (1976), Nitzan – Rubinstein (1981), Saari (1990), Debord (1992) y Van Newenhizen (1992), entre otros (véase Chebotarev – Shamis (1998)).

Aunque no se trata de una caracterización, queremos incorporar también en esta sección la "justificación parcial" de la regla de Borda proporcionada por Black (1976). Entendemos este calificativo en el sentido de que Black, al coincidir con las posiciones electorales de Condorcet y ser estas divergentes de las de Borda, nunca podría justificar totalmente la regla de propuesta por éste último, que puede proporcionar un ganador distinto del de Condorcet[24] (aquel que vence una a una a todas las demás alternativas por mayoría simple).

El razonamiento de Black (1976) retoma un enfoque utilitarista ya apuntado en Black (1958) y se basa, además de en la definición clásica de la cuenta de Borda, en un segundo

---

$(s_1 \geq s_2 \geq \cdots \geq s_n \geq 0,\ s_1 > s_n)$. De este modo, la regla de Borda no es sino un caso particular de regla de puntuación con pesos $s_i = n - i$. A su vez, la regla de Borda así definida es equivalente a cualquier regla de puntuación con pesos en progresión aritmética de diferencia común positiva, esto es, tales que $s_i - s_{i+1} = d > 0$, $i = 1, \ldots, n-1$. Nótese, así mismo, que la regla de pluralidad corresponde en este contexto a una regla con puntuaciones $s_1 = 1, \quad s_2 = \cdots = s_n = 0$.

[21] Ahora bien, un enfoque utilitarista-cardinal de las puntuaciones, que ya aparecía tácitamente en Borda [1770](1784), ha sido retomado por Black (1958 y 1976), Sugden (1981, pp. 140–145) y Marchant (2000), en este último caso con un tratamiento difuso (*fuzzy*) de las preferencias de los agentes.

[22] Sobre la polémica Borda – Condorcet y una vía de aproximación mediante sistemas de votación híbridos, véase Martínez Panero (en prensa).

[23] Esta es una propiedad de tipo arquimediano que indica que si se replican las preferencias de una coalición de votantes un número suficiente de veces, entonces esta coalición (junto con sus clones) deviene ganadora.

[24] Por tal razón se dice que la regla de Borda no es *Condorcet consistente*.

conteo individual que asigna a cada alternativa el número de alternativas a las que derrota menos el número de aquellas por las que es derrotada. Black (1976) demuestra entonces que existe una transformación afín entre ambos contadores (por lo que son equivalentes en el sentido de que ambos proporcionan la misma ordenación social) y que la suma de las puntuaciones emitidas por cada agente mediante el segundo contador es nula[25]. Con tales prerrequisitos, Black se decanta por el uso del segundo contador (por motivos más bien estéticos, de elegancia en las demostraciones aportadas, si se tiene en cuenta a Coughlin (1980)) y justifica entonces el uso de este segundo contador interpretándolo en términos mayoritarios. Además, argumenta que permite compensación[26], tiene en cuenta toda la información contenida en las preferencias mostradas por los agentes, proporciona un orden social transitivo[27], y selecciona como ganadora la alternativa que, en media[28], ocupa mejores posiciones en las preferencias de los agentes (algo que, como ya hemos comentado, ya aparecía explícitamente en Black (1958)).

Señala también Black, como ya habían hecho en el pasado Condorcet, Laplace y Daunou, que el voto estratégico puede arruinar la implementación efectiva y utilidad de la regla de Borda. Ahora bien, a partir del teorema de Gibbard – Sattterhwaite, se puede afirmar que tal susceptibilidad de manipulación no es exclusiva suya, sino común a todos los procedimientos de votación (véase Taylor (2005, pp. 60–69). Por este motivo, aunque la manipulabilidad, como también la inconsistencia de Condorcet del método de Borda (las dos principales críticas de Black) son incuestionables, éstas se deben analizar comparativamente con otros procedimientos, para lo que nos remitimos al análisis de Saari (Sección 5) que, como veremos, coloca a la regla de Borda en una posición de privilegio.

Cook *et al.*, Farkas – Nitzan: métricas sobre preferencias y metodología DEA

Cook – Seiford (1982) probaron, a partir de Kendall (1962), que la regla de Borda minimiza para la métrica 2 el desacuerdo total agregado (sobre las métricas $p,\ 1 \le p \le \infty$, véanse, Cook – Kress – Seiford (1996) y González Pachón – Romero (1999)). En la misma línea, en la que late la idea de consenso (*distance-consensus approach*), Farkas – Nitzan (1979) también demostraron, a su vez, que el método de Borda coincide con la regla de cercanía a la unanimidad (*closeness to unanimity*) introducida por los citados autores en términos paretianos sugeridos por Sen (1977). Ulteriores desarrollos de esta línea de investigación deben recoger las aportaciones de González Pachón – Romero (1999) utilizando programación por metas, y de Marchant (2001) siguiendo los pasos de Farkas – Nitzan (1979) en un contexto que incluye torneos, órdenes débiles, semiórdenes, relaciones de preferencia difusa, etc.

Cabe así mismo señalar que la metodología DEA (Análisis Envolvente de Datos) ha proporcionado una atractiva manera de afrontar el problema de la asignación de pesos. A este respecto Cook – Kress (1990) han probado que un caso particular de lo que denominan

---

[25] Véase también Coughlin (1980). En Martínez Panero (2003) se demuestran ambas propiedades de forma inmediata (utilizando el axioma de reciprocidad) para reglas de Borda difusas que extienden los casos clásicos formulados en Black (1976).

[26] Como ya señalamos, método "de compensación y suma" fue el nombre dado por Morales (1797) a la regla de Borda.

[27] Estas dos últimas características diferencian radicalmente la bondad del método de Borda con las "patologías" que pueden presentar los métodos de pluralidad, que sólo toma en cuenta las mejores alternativas, y de mayoría simple, que puede producir ciclos en el agregado (paradoja de Condorcet).

[28] Sobre los argumentos de tipo media en conexión con la regla de Borda, véanse también Straffin Jr (1980) y Mueller (1979).

función de discriminación de intensidad reproduce exactamente el modelos de Borda y Kendall. Por otro lado, también basándose en la metodología citada, Contreras – Hinojosa – Mármol (2005) han propuesto un método alternativo a las técnicas de consenso apuntadas anteriormente, evaluando las alternativas mediante pesos flexibles de manera que se obtenga una ponderación óptima para cada una de ellas. Y de nuevo, en su análisis, la regla de Borda aparece involucrada en la solución que se obtiene al resolver el problema de programación lineal que supone tal optimización.

Saari: geometría del voto y "optimalidad" de la regla de Borda

Saari (1995, p.20) se arroga el mérito de haber dicho la última palabra sobre el tema que nos ocupa al afirmar rotundamente que la cuestión de la particular elección de pesos de la regla de Borda, que había permanecido abierta durante los dos últimos siglos, ha sido por fin contestada por él mediante su "geometría del voto".

Desde tiempos de Borda [1770](1784) es bien conocido el hecho de que toda la problemática electoral surge cuando concurren 3 o más alternativas[29]. Por esta razón. de modo paradigmático (para mantener el aparato matemático a un nivel razonable), Saari (1994, 1995) ha estudiado exhaustivamente el voto ponderado en el caso de 3 alternativas[30], situándolas como vértices de un triángulo de manera que las 6 zonas en las que éste queda descompuesto después de trazar sus medianas corresponden a una de las 6 posibles ordenaciones de preferencias según la lejanía de un punto interior genérico de las mismas a los vértices. (Véase también Nurmi (1999)).

Mediante un análisis en el que se incide en la simetría intrínseca de la regla de Borda, Saari (1994, p.14) demuestra que tal método es el que minimiza la manipulabilidad de entre todos los sistemas de voto posicional, así como las posibles paradojas en las que pudieran incurrir tales procedimientos[31]. Por otro lado, Van Newenhizen (1992) y Saari (2000) han probado que este método es la regla de puntuación menos susceptible de vulnerar el criterio de Condorcet, o lo que es lo mismo, con mayor *eficiencia de Condorcet*[32]. Además, Saari – Barney (2003), incidiendo en un tema introducido en Saari (1995, p.153), también han probado que la simetría de la regla de Borda hace a este método él único entre los posicionales inmune a la paradoja de la reversión (*reversal bias*): para cualquier otra regla de puntuación existen situaciones no triviales[33] en las que, a pesar de invertirse las preferencias

[29] De hecho, problemas de agenda (orden de presentación de alternativas para su voto por parejas) con más de dos opciones ya fueron tenidos en cuenta estratégicamente por Plinio el Joven en el contexto del senado romano (véanse Farquharson (1969), Riker (1986, pp. 78–88) y McLean – Urken (1995, pp. 14–16 y 67–70)).

[30] En el caso de 3 alternativas y una vez realizado un proceso de normalización, un procedimiento de voto ponderado (o si se quiere, una regla de puntuación) viene dado por un vector $w_s = (1, s, 0)$ donde 1 es el peso asignado a la mejor alternativa, $s \in [0,1]$ es el que obtiene la segunda alternativa y 0 la tercera. Los valores extremos alcanzables por *s*, 0 y 1, corresponden a los procedimientos de pluralidad (cada agente selecciona la mejor alternativa) y antipluralidad (cada agente indica su peor alternativa). El caso intermedio $s = 1/2$ corresponde justamente a la regla de Borda.

[31] "[The Borda count] is the unique method to minimize the number and kind of [voting] paradoxes, to minimize the likelihood of a paradox, [and] to minimize the likelihood that a small group can succesfully manipulate the outcome".

[32] La eficiencia de Condorcet se define como la probabilidad condicionada de que un procedimiento seleccione el ganador de Condorcet, supuesto que tal ganador exista.

[33] Una situación trivial sería, por ejemplo, aquella en la que los órdenes lineales individuales fuesen todas las permutaciones circulares de uno dado, lo que provocaría un empate colectivo (o, desde un punto de vista no posicional, la aparición de un ciclo.

individuales, no hay cambio en la elección social resultante respecto de las situaciones originales (sobre la excepcionalidad de la regla de Borda entre los métodos posicionales, véase también Nurmi (2004 y 2005)).

Ahora bien, la militancia de Saari en pro de la regla de Borda, que considera óptima (Saari (1995, p. 19), no se restringe únicamente a su *status* entre los sistemas de voto posicional. Saari ha sostenido (y sostiene) polémicas intelectuales (que nos hacen recordar el enfrentamiento Borda *versus* Codorcet) con Brams y Fishburn (introductores y defensores del método de voto aprobatorio) y últimamente con Risse (que propugna métodos Condorcet consistentes, tales como la regla de Kemeny, en perjuicio de la regla de Borda). Sin embargo, a pesar de la profusión de trabajos en distintos frentes que avalan la implementabilidad de la regla de Borda por parte de Saari, está aún lejos de verificarse el desideratum de Young (1997) sobre su puesta en práctica[34].

Conclusiones

En el presente trabajo se ha presentado una panorámica de las distintas justificaciones a la presunta arbitrariedad o imposición de las ponderaciones que aparecen en la regla de Borda. Como hemos comentado, ya Arrow señaló que éste es un tema que viene repitiéndose de tiempo en tiempo y, efectivamente, las distintas argumentaciones probabilístico-estadísticas, axiomáticas, métricas y basadas en la metodología DEA, así como las geométricas que hemos expuesto parecen avalar tal aserto. Ahora bien, a partir de la variada gama de técnicas y métodos empleados que hemos reseñado podemos hacer la siguiente valoración: si la importancia de un problema científico depende en buena medida de las disciplinas que se generan o desarrollan para resolverlo y que son, a su vez, susceptibles de ser empleadas en otros contextos, el problema tratado en el presente trabajo, lejos de ser una cuestión aislada, ha resultado sumamente fértil, ya que, si bien su tratamiento no ha generado *per se* las técnicas citadas, el análisis implicado en su resolución ha coadyuvado incuestionablemente en el desarrollo de las mismas.

Para concluir, queremos indicar una posible vía que, en cierta forma, elude el problema planteado, ya que no lo encara como tal, sino en un contexto más amplio. Tal posibilidad consiste en diseñar reglas de Borda más generales, en las que las que, aún manteniéndose su filosofía de agregación, se permita a los agentes manifestar sus preferencias de forma matizada (gradual o lingüística), y no meramente mediante las preferencias taxativas que son la base informacional de la regla de Borda "clásica". De esta manera se generarían puntuaciones Borda que serían, según la forma de capturar lo más fielmente posible tales matices de información, números reales, intervalos, números difusos trapeciales o triangulares, dependiendo de la representación utilizada. Este es el tratamiento seguido en Marchant (1996a,1996b) García Lapresta – Martínez Panero (2000, 2002), García Lapresta – Llamazares – Martínez Panero (2005 y 2006) y García Lapresta – Martínez Panero – Meneses (en prensa).

---

[34] Young (1997, p. 200): "I predict that the time will come when […] Borda's rule will be considered a standard tool for legislative and committee decision making".

Agradecimientos

El autor agradece a Jesús Basulto, de la Universidad de Sevilla, que le facilitara facsímiles de las obras de Condorcet, Laplace y Todhunter utilizadas en la primera y segunda secciones del trabajo. Así mismo, reconoce el apoyo financiero proporcionado por la Junta de Castilla y León (Consejería de Educación y Cultura, Proyecto VA040A05), el Ministerio de Educación y Ciencia (Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica, Proyecto SEJ2006-024267/ECON) y FEDER.

Bibliografía

ARROW, K.J. (1974): *Elección Social y Valores Individuales*. Instituto de Estudios Fiscales, Madrid. [Traducción de Eusebio Aparicio Auñón de la segunda edición, corregida, en inglés, 1963: *Social Choice and Individual Values*. Introducción de Andreu Mas Colell. Primera edición, 1951].

BLACK, D. (1958): THE THEORY OF COMMITTEES AND ELECTIONS. KLUWER ACADEMIC Publishers, Boston.

BLACK, D. (1976): “Partial justification of the Borda count”. *Public Choice* 28, pp. 1–16.

BORDA, J.C. de (1784): *Mémoire sur les élections au scrutin*. Historie de l’Academie Royale des Sciences, París. [Se reproduce, en inglés, en Grazia (1953) y en McLean – Urken (1995, pp. 81–89)].

CONDORCET, J.A.M.N.C., marqués de (1785): *Essai sur l’Application de l’Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. L’Imprimierie Royale, París. [Se reproducen fragmentos escogidos, en inglés, en McLean – Urken (1995, pp. 91–112)].

CONTRERAS, I. – HINOJOSA, M.A. – MÁRMOL, A.M. (2005): “A class of flexible weight indices for ranking alternatives”. *IMA Journal of Management Mathematics* 16, pp. 71–85.

COOK, W.D. – Kress, M. (1990): “A data envelopment model for aggregating preference rankings”. *Management Science* 36 (11), pp. 1302–1310.

COOK, W.D.. – Kress, M. – Seiford, L.M. (1996): “A general framework for distanced-based consensus in ordinal ranking models”. *European Journal of Operational Research* 96, pp. 392–397.

COOK, W.D.. – Seiford, L.M. (1982): “On the Borda – Kendall consensus method for priority ranking problems”. *Management Science* 28, pp. 621–637.

COUGHLIN (1979): “A direct characterization of Black’s first Borda count”. *Economics Letters* 4, pp. 131–133.

CHEBOTAREV, P.Y. – Shamis, E. (1998): “Characterizations of scoring methods for preference aggregation”. *Annals of Operatons Research* 80, pp. 299–332.

DAUNOU, P.C.F. (1803): *Mémoire sur les Élections au Scrutin*. Baudouin, imprimeur de l’Institut National. París. [Se reproduce, en inglés, en McLean – Urken (1995, pp. 237–276)].

DEBORD, B. (1992): “An axiomatic characterization of Borda’s *k*-choice function”. *Social Choice and Welfare* 9, pp. 337–343.

FARKAS, D. – Nitzan, S. (1979): “The Borda rule and Pareto stability: a comment”. *Econometrica* 49, pp. 1305–1306.

Farquharson, R. (1969): *Theory of Voting*. Blackwell, Oxford.

FINE, B. – FINE, K. (1974): “Social choice and individual ranking I”. *Review of Economic Studies* 41, pp. 459–475.

FISHBURN, P.C. – Gehrlein, W.V. (1976): “Borda’s rule, positional voting and Condorcet’s simple majority principle”. *Public Choice* 28, pp. 79–88.

GARCÍA LAPRESTA, J.L. – LLAMAZARES, B. – MARTÍNEZ PANERO, M. (2005): “Reglas de Borda lingüísticas: análisis de sus propiedades”. En E. Herrera Viedma (ed.): *Procesos de Toma de Decisiones. Modelado y Agregación de Preferencias.* Universidad de Granada, Granada, pp. 69–78.

GARCÍA LAPRESTA, J.L. – LLAMAZARES, B. – MARTÍNEZ PANERO, M. (2006): “Linguistic matrix aggregator operators: Extensions of the Borda rule”. En B. Reusch (ed.): *Computacional Intelligence, Theory and Applications*. Springer Verlag, Berlín, pp. 561–576.

GARCÍA LAPRESTA, J.L. – MARTÍNEZ PANERO, M. (2000): “Análisis de algunos sistemas de votación a partir de la obra del ilustrado José Isidoro Morales”. *Hacienda Pública Española* 154, pp. 93–103.

GARCÍA LAPRESTA, J.L. – MARTÍNEZ PANERO, M.. (2002): “Borda count versus approval voting: A fuzzy approach”. *Public Choice* 112, pp. 167–184.

GARCÍA LAPRESTA, J.L. – MARTÍNEZ PANERO, M. – MENESES, L.C. (en prensa): “Defining the Borda count in a linguistic decisión making context. *Information Sciences*.

GÄRDENFORS, P. (1973): “Positionalist voting functions”. *Theory and Decision* 4, pp. 1–24.

GOODMAN, L.A. – Markowitz, H. (1952): “Social welfare functions based on individual rankings”. *American Journal of Sociology* 58(3), pp. 257–262.

GONZÁLEZ PACHÓN, J. – ROMERO, C. (1999): “Distance-based consensus methods: A goal programming approach”. *Omega* 27, pp. 341–347.

GRAZIA, A. DE (1953): “Mathematical derivation of an election system”. *Isis* 44, pp. 42–51.

HANSSON, B. – SAHLQUIST, H. (1976): “A proof technique for social choice with variable electorate”. *Journal of Economic Theory* 13, pp. 193–300.

HECKELMAN, J.C. (2003): “Probabilistic Borda rule voting”. *Social Choice and Welfare* 21, pp. 455–468.

INTRILIGATOR, M. (1979): “A probabilistic model of Social Choice”. *Review of Economic Studies* 40, pp. 553–560.

Kendall, M. (1962): *Rank Correlation Methods*. Hafner, Nueva York.

KENDALL, M. – STUART, A. (1977): *The Advanced Theory of Statistics. 1, Distribution Theory.* Charles Griffin, Londres.

LAPLACE, P.S. DE [1814] (1987): *Ensayo Filosófico sobre las Probabilidades.* Alianza Editorial, Madrid. Traducción del francés de Pilar Castrillo. [Hemos manejado también la edición de Alfredo B. Besio y José Banfi, Espasa Calpe, Buenos Aires, 1947].

LÓPEZ DE PEÑALVER, J. (1992): *Escritos*. Instituto de Cooperación Iberoamericana – Quinto Centenario – Antoni Bosch, editor – Instituto de Estudios Fiscales, Madrid. [Edición de Ernest Lluch de la obra original, 1799].

MARCHANT, T (1996a): *Agrégation de relations valuées par la méthode de Borda en vue d'un rangement. Considérations axiomatiques.* Ph. D. Thesis, Université Libre de Bruxelles, Bruselas.

MARCHANT, T (1996b): "Valued relations aggregation with the Borda method". *Journal of Multi-Criteria Decision Analysis* 5, pp. 127–132.

MARCHANT, T. (2000): "Does the Borda rule provide more than a ranking?" *Social Choice and Welfare* 17, pp. 381–391.

MARCHANT, T. (2001): "The Borda rule and Pareto stability: A further comment". *Fuzzy Sets and Systems* 120, pp. 423–428.

Martínez Panero, M. (2004): *Generalizaciones y Extensiones de la Regla de Votación de Borda.* Tesis doctoral, Universidad de Valladolid, Valladolid.

MARTÍNEZ PANERO, M. (en prensa): "Métodos híbridos de votación bajo preferencias usuales y difusas". *Anales de Estudios Económicos y Empresariales,* XVII.

MARTÍNEZ PANERO, M. – GARCÍA LAPRESTA, J.L. (2003): *José Isidoro Morales, Precursor Ilustrado de la Teoría de la Elección Social: Edición Facsímil de la Memoria Matemática sobre el Cálculo de la Opinion en las Elecciones (1797) y Apéndice (1805).* Secretariado de Publicaciones e Intercambio Editorial. Universidad de Valladolid, Valladolid. Prólogo de Salvador Barberà.

MASSÓ, J. (2003): Reseña de Martínez Panero – García Lapresta (2003). *La Gaceta de la Real Sociedad Matemática Española* 8, pp. 633–636.

MCLEAN, I. – Urken, A.B. (eds.) (1995): *Classics of Social Choice*. The University of Michigan Press, Ann Arbor.

MORALES, J.I. (1797): *Memoria Matemática sobre el Cálculo de la Opinion en las Elecciones*. Imprenta Real, Madrid.

MORALES, J.I. (1805): *Apéndice á la Memoria Matemática sobre el Cálculo de la Opinion en las Elecciones*. Imprenta de Sancha, Madrid.

MUELLER, D.C. (1979): *Public Choice.* Cambridge University Press, Londres. [Existe traducción al castellano de Juan Carlos Zapatero en Alianza Editorial, Madrid, 1984].

NITZAN, S. – RUBINSTEIN, A. (1981): "A further characterization of Borda ranking method". *Public Choice* 36, pp. 153–158.

NURMI, H. (1999): *Voting Paradoxes and How to Deal with Them*. Springer-Verlag, Berlín-Heidelberg.

NURMI, H. (2004): "A comparison of some distance-based choice rules in ranking environments". *Theory and Decision* 57, pp. 5–24.

NURMI, H. (2005): "A responsive system". *Economics of Governance* 6, pp. 63–74.

RIKER, W.H. (1986): *The Art of Political Manipulation.* Yale University Press, New Haven.

SAARI, D.G. (1990): "The Borda dictionary". *Social Choice and Welfare* 7, pp. 279–317.

SAARI, D.G.. (1994): *Geometry of Voting*. Springer–Verlag, Berlín.

SAARI, D.G. (1995): *Basic Geometry of Voting*. Springer–Verlag, Berlín.

SAARI, D.G. (2000): "Mathematical structure of voting paradoxes. II. Positional voting". Economic Theory 15, pp. 55–102.

SAARI, D.G. – Barney, S. (2003): "Consequences of reversing preferences". *Mathematical Intelligencer* 25(4), pp. 17–31.

SEN, A.K. (1977): "Social Choice Theory: A re-examination". *Econometrica* 45, pp. 53–89.

STEIN, W.E. – MIZZI, P.J. – PFAFFENBERGER, R.C. (1994): "A stochastic dominance analysis of ranked voting systems with scoring". *European Journal of Operational Research* 74, pp. 78–85.

Straffin Jr., P.D. (1980): *Topics in the Theory of Voting*. Birkhäuser, Boston.

SUGDEN, R. (1981): *The Political Economy of Public Choice: An Introduction to Welfare Economics*. Martin Robertson, Oxford.

TANGIAN, A.S. (2000): "Unlikelihood of Condorcet's paradox in a large society". *Social Choice and Welfare* 17, pp. 337–365.

TANGUIANE, A.S. (1991): *Aggregation and Representation of Preferences. Introduction to Mathematical Theory of Democracy*. Springer-Verlag, Berlín.

TAYLOR, A.D. (2005): *Social Choice and the Mathematics of Manipulation.* Cambridge University Press, Nueva York.

TODHUNTER, I. (1865): *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace.* Macmillan, Cambridge.

VAN NEWENHIZEN, J. (1992): "The Borda method is most likely to respect the Condorcet principle". *Economic Theory* 2, pp. 69–83.

YOUNG, H.P. (1974): "An axiomatization of Borda's rule". *Journal of Economic Theory* 9, pp. 43–52.

YOUNG, H.P. (1975): "Social Choice scoring functions". *SIAM Journal on Applied Mathematics* 28, pp. 824–838.

YOUNG, H.P. (1997): "Group choice and individual judgements", en Mueller, D. (ed.): *Perspectives on Public Choice: A Handbook*. Cambridge University Press, Cambridge, pp. 181–200.

Capítulo 16

Kolmogorov y sus aportaciones fundamentales

CRISTINA SÁNCHEZ FIGUEROA
PEDRO CORTIÑAS VÁZQUEZ
IÑIGO TEJERA MARTÍN
Universidad Nacional de Educación a Distancia (Uned)

Introducción

La Estadística es una ciencia con tanta antigüedad que su proceso de desarrollo hace, que por sí misma, sea auxiliar de otras ciencias como: la medicina, la ingeniería, la sociología, etc, y está presente en las más diversas esferas de la vida cotidiana, como se observa en el deporte que las marcas de un mismo atleta son diferentes en distintas repeticiones del mismo ejercicio. La Estadística actúa como disciplina que proporciona técnicas precisas para obtener información, (recogida y descripción de datos) y por otra parte proporciona métodos para el análisis de la misma. Los métodos y técnicas estadísticas ayudan a la realización de múltiples tareas en las organizaciones productivas o sociales, tanto públicas como privadas; así la ausencia de ésta conllevaría a un caos generalizado no existiendo en muchos casos información vital a la hora de tomar decisiones en tiempos de incertidumbre. Del mismo modo, es la base para la realización de estudios e investigaciones que permiten la mejora, no solo, de los procesos de producción, de bienes y de servicios, sino también de las decisiones tomadas en las empresas u organizaciones de los más diversos ámbitos.

Aunque los comienzos de la estadística pueden ser hallados en el antiguo Egipto con los recuentos de población o de riqueza, la estadística podemos decir que ha existido desde el inicio de las civilizaciones, si bien, utilizando técnicas que hoy consideramos elementales pero que tienen toda su importancia por el momento en el que se promueven. Durante muchos años las grandes operaciones estadísticas fueron la realización de censos o los recuentos de defunciones, no obstante la estadística que conocemos hoy en día debe gran parte de su elaboración a los trabajos matemáticos de aquellos hombres que desarrollaron la teoría de la probabilidad. Durante el siglo XVII y principios del XVIII, matemáticos como Bernoulli, Moivre, Bayes, Lagrange y Laplace desarrollaron la teoría de probabilidades. Pese a que durante cierto tiempo la teoría de la probabilidad limitó su aplicación a los juegos de azar, no será hasta el siglo XIX cuando comienza a aplicarse a los grandes problemas científicos.

Será en el año 1933 cuando el célebre matemático estadístico Kolmogorov se propuso construir una teoría de la probabilidad totalmente rigurosa, formuló un sistema de axiomas para la teoría de la probabilidad, basado en la teoría de conjuntos y en la teoría de la medida, desarrollada pocos años antes por Lebesgue, Borel y Frechet entre otros. La construcción axiomática de la teoría de la probabilidad, procede de las propiedades fundamentales de la probabilidad observadas en los ejemplos que ilustran la definición clásica y frecuentista. La definición axiomática las incluye como casos particulares y supera las carencias de ambas. Así la probabilidad pudo desarrollarse como una teoría completamente lógica, al mismo tiempo que siguió permitiendo resolver los problemas aplicados.

Un científico tan impresionante como Kolmogorov recibió un amplio reconocimiento honorífico de muchos países, pues a lo largo de su trayectoria científica escribió muchos libros y más de 200 artículos sobre diversas disciplinas científicas como física, mecánica o biología. Pero será a los diversos campos de la matemática a los que dedique su mayor número de artículos, destacando entre sus aportaciones las que realiza a la Teoría de Conjuntos, Teoría de Integración, Teoría de la Información, Teoría de la Turbulencia, *Teoría de Probabilidad*, *Teoría de los Procesos Estocásticos, o Estadística Matemática* entre otras.

Andrei Nikolaevich Kolmogorov.

Es importante señalar que el trabajo científico llevado acabo durante toda su vida por el célebre matemático estadístico Andrei Nikolaevich Kolmogorov es un ejemplo de superación y de gran dedicación a la investigación aplicada (no solo en el campo de las matemáticas sino también en el de otras disciplinas), pues los años en los que se desarrolla están condicionados, entre otros acontecimientos, por las revoluciones sociales (ocasionadas por las condiciones de crisis económica vivida) y las dos guerras mundiales.

Sus padres no pudieron influir en su educación, su padre murió en 1918 durante la guerra civil luchando en las filas del ejército rojo y su madre murió durante el parto. Así, el pequeño Andrei creció en una de las propiedades de su abuelo materno, donde cuidaron de él sus tías maternas quienes se preocuparon en desarrollar su curiosidad por la naturaleza y su interés en los libros.

Las primeras clases las recibe en una pequeña escuela creada por una de sus tías, para una docena de niños de vecindad. En la escuela se editaba la revista “Spring Swallows” donde publicó sus descubrimientos iniciales y algunos problemas aritméticos. A los seis años escribió su primer descubrimiento, la suma de los primeros números impares consecutivos es siempre un cuadrado:

$$1 = 1^2 \qquad 1+3 = 2^2 \qquad 1+3+5 = 3^2 \qquad 1+3+5+7 = 4^2 \qquad ,etc.$$

A los siete años se traslada a Moscú con su tía y allí se matriculará en el Instituto Repman. Las clases seguían métodos no tradicionales, pero muy avanzados para su tiempo. En Repman contó con buenos profesores, destacó obteniendo premios en historia y matemáticas y comenzó su interés, más serio, por la investigación científica.

En 1920 comienza a estudiar en la Universidad Estatal de Moscú, en la Facultad de Física y Matemática. La teoría de conjuntos con Zhegalkin y la geometría proyectiva con Vlasov, fueron sus primeros cursos, luego se interesó en la teoría analítica de funciones enseñada por Nikolai Nikolaievitch Luzin. Su primer trabajo matemático surgió de un contraejemplo en una proposición geométrica elemental planteada por Luzin, así aunque Kolmogorov era tan sólo un estudiante comenzó a investigar y producir resultados de importancia internacional en este campo de la ciencia.

También asiste al seminario sobre series trigonométricas que dirigía el profesor Viacheslav Stepánov, en él resuelve el problema de la construcción de una serie de Fourier cuyos coeficientes tienden a cero mas lentamente de lo deseado, un problema en el cual Luzin estaba muy interesado. En el artículo resultante formula un resultado más completo sobre los valores de los coeficientes de Fourier. También obtuvo su más celebre resultado sobre series trigonométricas construyendo una serie de Fourier-Lebesgue que diverge en casi todos los puntos. Su interés por la teoría de series trigonométricas y teoría descriptiva de conjuntos, hizo que sus investigaciones le llevaran al análisis clásico (diferenciación, integración, teoría de la medida y lógica matemática). Kolmogorov analizó en profundidad las afirmaciones y nuevas construcciones de la integral, así vuelve a sorprender a la comunidad científica construyendo una función integrable cuya representación en serie trigonométrica divergía en todo punto.

En años sucesivos comienza su interés por la *teoría de la probabilidad*, así en 1925 tiene lugar otro momento importante en su carrera pues publica su primer artículo sobre *probabilidad.* Fue publicado conjuntamente con Aleksandr Yakovlevich Khinchin y contiene *"El teorema del las tres series"*, en él formula la convergencia de una serie, si y solo si, converge en probabilidad considerando para ello tres supuestos fundamentales. El interés por el azar y las cuestiones relacionadas con la probabilidad, son temas que marcarán el inicio de los posteriores trabajos científicos desarrollados a lo largo de toda su vida.

Kolmogorov termina su carrera en 1925, pero este mismo año decide iniciar sus estudios de postgrado, durante cuatro años más, en la Universidad Estatal de Moscú. Es importante señalar que entre 1922 y 1925, años en los que realiza su carrera, también desarrolla la labor de Maestro de Matemáticas y Física en la escuela modelo experimental de Moscú, este entusiasmo por la enseñanza lo acompañará a lo largo de toda su vida. Así y tras esta experiencia, en 1931 Kolmogorov fue promovido como profesor en la Universidad Estatal de Moscú. Será también en este mismo año cuando publique un artículo fundamental *"Métodos analíticos de la teoría de probabilidades"*, que será el inicio de lo que hoy conocemos como procesos de Markov.

En años sucesivos, siendo director del Instituto de Investigaciones en Mécanica y Matemática, publica su monografía sobre *"Fundamentos de la Teoría de Probabilidades"* cuyo nombre original en alemán es *"Grundbegriffe der Wahrscheinlichkeitsrechnung"* (1933). En ella reconstruyó la teoría de probabilidad de forma rigurosa, a partir de los axiomas fundamentales que hoy en día se reconocen y aceptan al estudiar la teoría probabilidad. Serán precisamente sus trabajos sobre la teoría de probabilidades y sus aplicaciones, los que le consolidan y le atribuyen gran relevancia científica a nivel mundial. Esto hace que entre 1938 y 1966 sea *Jefe de la Cátedra de Teoría de Probabilidades*; creada a instancia suya ante la necesidad de seguir investigando estadística matemática.

De igual forma plantea un nuevo enfoque de la teoría de la información, el concepto de complejidad que va más allá de los postulados que partían de una concepción estadística de la información. Así, Kolmogorov en el año 1963 expone la idea de proporcionar una base algorítmica a las nociones básicas de la teoría de la información y *la teoría de probabilidades.*

A pesar de que su apasionante vida científica giró entorno al desarrollo matemático, Kolmogorov también tuvo tiempo, no solo, para llevar a cabo su vida personal, que compartirá hasta su muerte con su amiga de la escuela Anna Dimitrieva Egórova con quien se casa en 1942, sino también para contactar con otros científicos no matemáticos quienes le instruirán sobre temas como la física o la biología. Serán estas aportaciones las que precisamente lleven a Kolmogorov a realizar el análisis de los problemas de las turbulencias, desde el punto de vista estocástico, y el estudio de los procesos ramificados. Así entre 1946-1949 es jefe del laboratorio de investigaciones sobre turbulencia en el Instituto de Geofísica Teórica.

Para Kolmogorov su vida siempre estuvo repleta de prosperidad que recibió tanto por el reconocimiento a sus logros personales, en la búsqueda de la verdad en el proceso de conocimiento, como de la dedicación a fines humanos, ya que participó con gran entusiasmo en la educación de los jóvenes a lo largo de su vida científica. Como reconocimiento a sus meritos didácticos, llega a ser miembro fundador de la Academia de Ciencias Pedagógicas de la URSS en 1966 y dedica los últimos veinte años de su vida a la obra pedagógica. El interés por la educación le llevó a tener estudiantes con un talento excepcional, con los que realizó una importante labor científica, entre ellos destaca Vladimir I. Arnold con quien resuelve el 13 problema de Hilbert.

Si analizamos la diversidad de trabajos y artículos, publicados a la largo de la vida de Kolmogorov, podemos tener una imagen incierta de un científico que analiza materias dispersas y sin ningún tipo de relación. Pero es todo lo contrario, su inquietud científica hace que se preocupe por alcanzar objetivos suficientemente ambiciosos en temas relacionados con la matemática, *teoría de probabilidades, estadística matemática*, teoría de sistemas y procesos estocásticos, no solo para su propia satisfacción personal sino también como una excelente contribución a la comunidad científica. Toda una vida dedicada a la investigación científica llevó a que recibiera las más altas distinciones del sistema soviético, y al mismo tiempo fuera elegido miembro honorario de las academias de ciencia y de las sociedades científicas más prestigiosas del mundo. Haciendo un repaso posemos enunciar las siete órdenes Lenin, el premio Stalin, el premio Lenin que le concedió el Estado Soviético por trabajos en la teoría de perturbaciones de sistemas dinámicos. Además también recibió el premio Chebyshev por su libro *"Distribuciones límite para sumas de variables estocásticas independientes"*, el premio internacional Fundación Balzan en Matemáticas, premio internacional Fundación Wolf por profundos y originales descubrimientos en el *"análisis de Fourier, teoría de probabilidades, teoría ergódica y sistemas dinámicos"*, y el premio Lobatchevski de la Academia de las Ciencias de la URSS en 1987. Este mismo año su cerebro dejó de pensar para siempre y sus trabajos quedaron para que fueran recopilados y publicados en años sucesivos .

Aportaciones a la Estadística.

Es fundamental señalar la gran aportación que realizó a la Estadística, como consecuencia de su inquietud científica.

> *"Andrei Nikolaievitch ya no está, pero la naturaleza guarda el eco de sus preguntas y los hombres conservan su mensaje de esperanza. La vida de todo científico completo es parte de un torrente inagotable que labra el curso de la humanidad"*[1].

---

[1] Rolando Rebolledo

- FUNDAMENTOS DE LA TEORÍA DE PROBABILIDADES

Uno de los aspectos más significativos de Kolmogorov es su gran interés por cuestiones relacionadas con el azar, esto le llevará a ser un gran precursor de la *teoría de la probabilidad*. Desde el inicio de su trabajo científico escribe diversos artículos sobre *probabilidad*, el primero en 1924 siendo aún estudiante de la universidad; en él desarrolla la convergencia de series cuyos términos dependen del azar. Pero el trabajo que le da relevancia a nivel mundial es *Fundamentos de la Teoría de Probabilidades, cuyo nombre original en alemán es "Grundbegriffe der Wahrscheinlichkeitsrechnung" (1933).* Las definiciones dadas hasta el momento eran demasiado intuitivas y con una elevada experiencia práctica, por lo que no eran muy apropiadas para un desarrollo teórico de la teoría de la probabilidad. Será a finales del siglo XIX y principios del siglo XX, cuando se pueda realizar una formalización más exhaustiva de los conceptos probabilísticos gracias al desarrollo de las teorías de conjuntos y de la medida de Georg Cantor, Émile Borel y Heri Lebesgue. Así Kolmogorov, poniendo en relación estas teorías construye *una teoría de la probabilidad* de forma rigurosa y basándose en unos axiomas fundamentales. Esta aproximación axiomática surge ante las limitaciones surgidas de las teorías:

Clásica: Esta definición clásica de probabilidad fue una de las primeras que se dieron y se atribuye a Laplace; En 1812 Laplace publica su famosa obra "*Theoríe Analytique des probabilités*", que contiene una exposición completa y sistemática de la teoría matemática de los juegos de azar. Laplace define la probabilidad de un suceso como el cociente entre el número de casos favorables y el número total de casos, siempre que todos sean igualmente posibles. Pese a la gran cantidad de aplicaciones de esta teoría (que se conoce con el nombre de probabilidad a priori) existen también voces críticas, pues presenta una serie de inconvenientes que hacen su aplicación limitada.

Frecuentista: La definición frecuentista considera la probabilidad como una frecuencia relativa ideal. Define la probabilidad como el límite de la sucesión de frecuencias relativas de un suceso al aumentar indefinidamente el número de realizaciones. Aunque es imposible llegar a este límite, pues no podemos repetir el experimento un número infinito de veces, si podemos repetirlo muchas veces y observar como las frecuencias relativas tienden a estabilizarse.

Tanto en la definición clásica como en la frecuentista la probabilidad se define como un cociente, siendo siempre el numerador menor o igual que el denominador porque corresponde al número de elementos de un subconjunto del conjunto cuyos elementos contamos en el denominador. Así estas teorías hacen imposible la formalización matemática de la asignación de un modelo matemático a la probabilidad, pero Kolmogorov tenía gran empeño en convertir el cálculo de probabilidades en una autentica teoría, totalmente unida a la práctica y formalizada al estilo que Hilber proponía para toda su matemática.

En el enfoque axiomático Kolmogorov define probabilidad a partir de las propiedades fundamentales que debe satisfacer la noción de probabilidad, observadas en los ejemplos que ilustran las definiciones clásica y frecuentista, y que tomadas como puntos de partida o axiomas permiten edificar sobre ellas toda la teoría. Así conceptos ampliamente utilizados en la jerga probabilística, como variable aleatoria y función de distribución, cobran precisión matemática.

Las probabilidades son valores de una función de conjunto, también conocida como medida de probabilidad, esta función asigna números reales a diversos subconjuntos de un espacio muestral:

Sea ω un espacio muestral y sea $\overline{A}$ una colección de subconjuntos del espacio muestral ω, es decir un conjunto de sucesos. Se define probabilidad como una aplicación $P:\overline{A}\rightarrow[0,1]$ (a cada $A\in\overline{A}$ le hace corresponder un $P(A)$) que cumple los siguientes axiomas:

Axioma 1. Si A es un elemento de $\overline{A}$, existe un número $P(A)\geq 0$, denominado probabilidad del suceso A.

Axioma 2. $P(\omega)=1$

Axioma 3. Dada una sucesión númerable de sucesos, $A_1,\cdots,A_j$, disjuntos dos a dos, $A_i\cap A_j=\emptyset$, se verifica que: $P\left(\bigcup\limits_{i=1}^{n}A_i\right)=\sum\limits_{i=1}^{n}P(A_i)$

Así la probabilidad pudo desarrollarse como una teoría completamente lógica después de un siglo donde había sido dejada de lado, fue ganando aceptación como una teoría con innumerables aplicaciones a ramas muy diferentes del conocimiento. Numerosos probabilistas de la época, después de analizar la obra de Kolmogorov, quedaron convencidos de que la *teoría de probabilidades* puede desarrollarse en términos de la teoría de la medida, de forma tan rigurosa como cualquier otra área de la matemática. Pero el verdadero significado de esta obra se puede apreciar mejor en la actualidad, pues no sólo contiene la construcción axiomática de la teoría, sino que desarrolla de forma clara y precisa una serie de conceptos probabilísticos que serán la base necesaria para la fundamentación de los *procesos estocásticos*.

- *TEORÍA DE LOS PROCESOS ESTOCÁSTICOS*

El año 1931 es de crucial importancia para la constitución de los procesos estocásticos como una rama ampliamente desarrollada de la teoría de probabilidades. Este año Kolmogorov publica su artículo *Métodos analíticos en la teoría de probabilidades,* que será el inicio de lo que hoy conocemos como procesos de Markov, pues Kolmogorov descubrió profundas relaciones entre los procesos markovianos y las ecuaciones diferenciales.

Los procesos estocásticos o procesos aleatorios constituyen una herramienta estadística que surge ante la necesidad de modelar el comportamiento de experimentos aleatorios que varían en el tiempo o dependen de alguna otra variable aleatoria determinista. Los procesos estocásticos se pueden clasificar atendiendo a dos aspectos: si el espacio de estados posibles de la variable aleatoria contiene valores discretos o continuos y si los valores del tiempo son discretos o continuos. Asimismo, las cadenas de Markov constituyen un proceso estocástico en el que los valores del tiempo son discretos y los estados posibles de la variable aleatoria contienen valores discretos, es decir, es una cadena estocástica de tiempo discreto. De esta manera Kolmogorov desarrolla la *ecuación de Chapman-Kolmogorov,* como una identidad que deben satisfacer las probabilidades de transición para cualquier proceso de Markov.

- *ESTADÍSTICA MATEMÁTICA.*

o La prueba de Kolmogorov-Smirnov:

Otra gran aportación de Kolmogorov es la demostración de que la función de distribución empírica se aproxima, tanto como se quiera, a la distribución de probabilidades de la variable aleatoria, siempre que la muestra que se tome sea suficientemente grande. La forma de estimar el error máximo que se comete en esta aproximación, hizo que formara parte de todos los textos de estadística matemática.

La prueba de Kolmogorov-Smirnov (K-S), nombre con el que se conoce, fue propuesta originalmente en los años 30 por Andrei Nikolaevich Kolmogorov y Nikolai Vasilyevich Smirnov (1933). Es una prueba no paramétrica de bondad de ajuste que sirve para contrastar la hipótesis nula de que la distribución de una variable se ajusta a una determinada distribución teórica de probabilidad. Su objetivo fundamenta es señalar si los datos provienen de una población que tiene la distribución teórica especificada. Es una extensión del Teorema de Glivenco-Cantelli (llamado también *Teorema Fundamental de la Estadística* por el papel que desarrollará dentro de la Inferencia Estadística) que demuestra que para muestras grandes la función de distribución empírica $F_n(x)$ converge en probabilidad a la función de distribución de la población $F(x)$. Kolmogorov y Smirnov (1933) establecen la medida más simple para calcular la diferencia que existe entre ambas funciones, por medio de la distancia máxima (medida en dirección vertical) entre las gráficas correspondientes a ambas funciones, esa medida viene dada por lo que hoy conocemos como *el estadístico de Kolmogorov y Smirnov:*

$$D_n = Máx_{-\infty<x<\infty}\left|F_n(x) - F(x)\right|$$

La distribución de este estadístico D_n no depende del modelo fijado para $F(x)$ fijado en la hipótesis nula, es decir, es la misma para cualquier función $F(x)$.

- Teoría de los estimadores insesgados:

En 1950 Andrei Nikolaevich Kolmogorov completó uno de sus trabajos más importante en Estadística con el título de *Estimadores Insesgados*, en él analiza sistemáticamente las propiedades de los estimadores insesgados y los diferentes métodos de construcción, por medio de estadísticos suficientes, y también describe el significado de las aplicaciones de estos estimadores en problemas de estadísticas de control y control de calidad en la industria. Para ello recurre al teorema de Rao-Blackwell (1945-1947) para calcular el estimador insesgado de la proporción.

Bibliografía

CASAS SÁNCHEZ JOSÉ M. (1997)*: Inferencia Estadística.* Editorial Centro de Estudios Ramón Areces S.A.

GIRÓN GONZÁLEZ-TORRE, FRANCISCO JAVIER: (2004): *Breve introducción a la obra de A. N. Kolmogorov (1903-1987).* Real Academia de Ciencias y Universidad de Malaga.

KOLMOGOROV A.N. (1991-1993), *Vol. II : Probability Theory and Mathematical Statistics.* Kluwer, Berlin.

KOLMOGOROV A.N. (1956): *Foundations of the theory of probability,* Chelsea, New York.

KOLMOGORV IN PERSPECTIVE (2000): *History of Mathematics,* Vol 20. American Math. Soc. Y London Math. Soc.

MARTÍN PLIEGO, F. JAVIER Y RUIZ-MAYA, L. (1997): *Estadística. I. Probabilidad.* Editorial AC.

RUIZ MUÑOZ, D.: *Manual de Estadística.* Universidad Pablo Olavide.

SÁNCHEZ FERNANDEZ,C.Y VALDÉSCASTRO,C. (2003): *Kolmogórov , El zar del azar.* Editorial NIVOLA, libros y ediciones S.L.

Direcciones de Internet.:

1. http://Kolmogorov.com
2. http://Thales.cica.es/rd/Recursos/rd97/Biografias/40-1-b-Kolmogorov.html

Capítulo 17

El Desarrollo de las estadísticas del sector pesquero durante los siglos XVIII y XIX

JUAN JOSÉ GARCÍA DEL HOYO
Universidad de Huelva

Introducción

El conocimiento preciso de la situación real del sector pesquero, como base fundamental del sistema de reclutamiento de la Real Armada a través de la "Matrícula de Mar", se manifestó como una necesidad apremiante a lo largo del siglo XVIII, debido a la expansión de la marina española y a las necesidades de ésta derivadas de los múltiples conflictos bélicos en los que se vio implicada, sobre todo durante la segunda mitad del siglo, la corona borbónica para defender el comercio atlántico y la integridad del imperio colonial. Si bien fue ésta la principal causa para el desarrollo de un sistema de recogida de información sobre el sector, no puede olvidarse el interés económico derivado de la propia actividad. El Reino de España, como Estado estrictamente católico, tenía una necesidad permanente de abastecimiento de pescado a causa de la elevada demanda derivada de los periodos de ayuno, abstinencia y vigilia marcados por el calendario religioso, según el cual casi en el 40% de los días estaba vedado el consumo de carne y el necesario aporte proteínico debía sustentarse, básicamente, en los productos pesqueros. No es casual, por tanto, que desde la edad media existiese una relevante industria salazonera que abastecía una elevada demanda, que desde principios del siglo XVII se veía sometida a la fuerte competencia de las producciones extranjeras, fundamentalmente del bacalao inglés y del arenque holandés. Esta pugna comercial puede ser considerada como la causa indirecta de algunos conflictos bélicos o, al menos, como una de las bazas que se jugaron en las mesas de negociaciones, como prueban la práctica expulsión de los pescadores españoles de los bancos bacaladeros de Terranova tras el Tratado de Utrecht en 1714, el cierre del abastecimiento de la sal portuguesa a los pescadores de las Provincia Unidas en 1580, tras la anexión española del Reino, la conquista por éstos de las salinas del norte de Brasil entre 1630 y 1654, las presiones británicas desde 1680 a dirigentes del litoral atlántico africano para que impidiesen a los pescadores andaluces y canarios la instalación de sus tradicionales factorías de salazones de cazón, o la guerra abierta entablada en 1652 por la

Inglaterra puritana de Oliver Cromwell y las Provincias Unidas por el control de las pesquerías de arenque del Mar del Norte. Y es que el mercado español de productos de la pesca era, como aún hoy lo es, una fruta muy apreciada por todas las potencias marítimas, dado que la existencia de un potente sector pesquero equivalía a poder disponer de una marinería experta para las armadas europeas y, por consiguiente, permitía a éstas incrementar su poderío naval y expandir sus imperios coloniales.

La respuesta española a los retos planteados no fue otra que la consolidación de la "Matrícula de Mar" para permitir la expansión de la armada así como para garantizar la independencia del abastecimiento de pescado. Pero había una tercera razón para la consolidación del sector pesquero. El comercio a gran escala de productos de la pesca implicaba el concurso de grandes cantidades de sal, dado que éste era prácticamente el único recurso existente para la conservación de estos productos. La Corona detentaba desde la Edad Media el monopolio de su abastecimiento, que constituía un "producto estancado", y ésta fijaba el precio a su conveniencia, estableciéndose, además, frecuentes recargos sobre éste para obtener recursos adicionales para la Hacienda Pública, tanto para financiar políticas concretas de infraestructuras – red de carreteras - como para obtener fondos con los que hacer frente a las necesidades derivadas de los conflictos bélicos de finales del XVIII.

Pero paralelamente a este proceso expansivo de la pesca – y de su control – impulsado desde la Corona a través de la Real Armada, ésta no había quedado fuera del debate económico mercantilista que durante el siglo XVII y el XVIII se proponía la reforma del Estado para su recuperación y consolidación. No obstante, son los primeros ilustrados los que dedican una mayor atención a la actividad generando, en memoriales producidos al amparo de las Sociedades Económicas o en contextos conflictivos motivados por la introducción de nuevas técnicas extractivas, una gran cantidad de documentos que, desde una perspectiva económica, describen la situación de determinadas pesquerías, recopilando abundantes datos estadísticos e, incluso, ayudándose de series seculares, elaborar teorías concretas que hoy podríamos fácilmente considerar como el origen del concepto de pesca sostenible[1]. En muchos casos, como es el del Comisario Sañez Reguart, coincide el papel de inspector de matrículas con el de científico ilustrado, el de legislador y el de mediador en conflictos pesqueros.

Todas estas corrientes, que confluyen en la segunda mitad del siglo XVIII, constituyen la base sobre la que se afianzará el desarrollo inicial de las Estadísticas Pesqueras en España.

La Matrícula de mar y los primeros registros normalizados

La Matrícula de Mar, como institución creada para garantizar el reclutamiento de las tripulaciones de la Real Armada se instituye en España, de forma definitiva, mediante la denominada "Ordenanza del Infante Almirante" promulgada el 18 de octubre 1737[2], cuyo

---

[1] El concepto de pesca sostenible, tan en boga durante el siglo XX, fue esbozado por Martín de Sarmiento en 1757 en una Memoria sobre las almadrabas dirigida al Duque de Medina Sidonia. Véase García del Hoyo (2006), pág. 979-980.

[2] No obstante, debemos resaltar que existen precedentes de dicha institución en las "*Ordenanzas para las Armadas del Mar Océano*" que promulgó Felipe III en 1606, renovadas posteriormente por la Real Cédula de 10 de octubre de 1625 de Felipe IV. Éstas apenas contenían privilegios concretos excepto los de estar exentos de cargos en los concejos municipales o la preferencia, una vez jubilados y siendo armadores, de su embarcación sobre otras para fletes de la Corona. Nada relativo al privilegio exclusivo de la pesca, que es el que facilitó el éxito en la siguiente etapa. Lógicamente, se produjo una fuerte resistencia a estas normas que prácticamente invalidó sus resultados, por lo que es con la entrada de la dinastía borbónica cuando mediante el R. O. de 28 de

modelo se encuentra en el modelo francés implantado por Luis XIV. De hecho, desde principios del siglo XVIII se habían dado pasos en el sentido de proporcionar un instrumento para revitalizar la marina a través del incremento de las posibilidades de reclutamiento, buscando un sistema que permitiese que las levas de la Armada contaran con marinería experta para tripular las embarcaciones, lo que no podría lograrse sino a través del fomento de la pesca y la expansión de las actividades comerciales.

La situación de la Real Armada y de la flota mercante española a principios del XVIII era dramática; el tráfico comercial con las indias e incluso la protección de las propias flotas había quedado en manos de los aliados de Felipe V, al no contar la corona con navíos suficientes para garantizarlo. No es extraño, por esta razón, que la principal preocupación de nuestros principales mercantilistas sea la revitalización del comercio mediante la expansión de la flota nacional, para lo que creían necesario, acertadamente, fomentar todas las actividades marítimas y pesqueras que, debido al corso berberisco, a barreras de carácter institucional e impositiva, a la fuerte competencia que en el mercado interno mantenían las producciones holandesas y británicas y, en definitiva, al reducido atractivo que éstas ejercían, habían menguado durante toda la segunda mitad del XVII y la primera década del XVIII. Por tanto, la expansión de la pesca como medio de garantizar una marinería capaz y experta, era una de las bases sobre las que recaía el discurso mercantilista. De hecho, Gerónimo de Uztariz, nuestro principal mercantilista y el único economista español citado por Adam Smith, realiza en 1724 un primer esbozo del sistema que debería instituirse. "…y habiendo más pescadores en los Mares de España,…, como he apuntado, se aumentará también considerablemente el número de la Marinería, por la facilidad con que la gente moza y la adulta se inclinan a este genero de servicio de Mar,…, y se debe creer que estando acostumbrados a el, no tendrán dificultad o repugnancia en alistarse a servir en la Armada y en viajes largos"[3]. Y una vez fomentada la marinería debería establecerse algún sistema que permitiese el control de ésta y garantizase el reclutamiento: "…en habiéndose aumentado considerablemente la gente de mar, como se debe esperar, practicándose los medios que se proponen, será muy conveniente, que por Comisarios que residan en las mismas Provincias Marítimas, se formen listas de los Marineros que hubiere en cada una de ellas, con sus filiaciones, y nota de sus edades, tiempo y parajes que han servido en la Mar, y las demás circunstancias que se deben tener presentes, y que se practican en otros Reinos, y particularmente en Francia…".[4] Por tanto, la Matrícula de Mar no fue fruto exclusivamente, como algunos apuntan, de un cuerpo de oficiales de la Armada encabezado por José Patiño, sino que, siguiendo el modelo francés, es una propuesta del mercantilismo como remedio a la ruina del comercio.

En síntesis, el sistema garantizaba la libertad y exención del sorteo de Quintas de "toda la Gente de Mar, que quisiere matricularse, y alistarse para servicio de mis Navíos, y los Carpinteros de Ribera, y Calafates, que así mismo se matricularen para construirlos, carenarlos, y ponerlos en estado de navegar", exención que se extendía a la obligatoriedad de alojamiento de oficiales y soldados en tránsito, así como estar liberados de cualquier carga concejil o estar exentos absolutamente de la Jurisdicción Ordinaria, situándose bajo la

enero de 1717 y la R. O. de 29 de agosto de 1726 cuando se recupera dicha institución, tomando carta de naturaleza en la Ordenanza mencionada.

[3] Véase, Uztariz, J., (1724), pág. 229.

[4] Sus aportaciones sobre lo que después se llamaría "*Matrícula de Mar*" no se limitan a estas recomendaciones, sino que incluso propone la rotación de periodos, recogida posteriormente en las Ordenanzas de 1751 y la concesión de privilegios fiscales y de otras cargas "*…y será muy justo también que cuando hubieren servido dos o tres años a su Majestad, gocen algunos privilegios, como es la exención de alojamientos militares y de las cargas concejiles, según lo practicó también el Rey Luis XIV, siendo gracias que no agravan a la Real Hacienda*" [Ídem, pág. 230].

Jurisdicción o Fuero de Marina.[5] Pero el privilegio de mayor interés reside, sobre todo, en la exclusión de todas las actividades pesqueras – excepto "la pesca de Vara ó Caña, y la de los Esparabeles, ó artes de pescar, de que puedan usar desde tierra, sin valerse de embarcaciones" – a todos los que no estuvieses matriculados[6]. A cambio, los matriculados tendrían que estar disponibles para el servicio en la Real Armada hasta alcanzar la edad de jubilación de éste - que se alcanzaba tras haber servido treinta años seguidos o por haber alcanzado la edad de sesenta años – pudiendo ser requerido para el servicio activo durante un año de cada cuatro o más de forma excepcional "en tiempos de guerra"[7]. No obstante, existía la posibilidad de que para hacer frente a estas circunstancias excepcionales sin que se dañase de forma irremediable al sector pesquero, se admitiesen en los buques de la armada hasta una tercera parte de la tripulación de no matriculados.

Lógicamente, para que todo el sistema funcionase correctamente era necesario establecer una organización territorial y orgánica que permitiese el control efectivo de la matrícula y, lo que era más importante, la defensa de sus privilegios frente a las pretensiones de los Concejos y de los Señoríos territoriales eclesiásticos y civiles. Para ello, la Ordenanza de 1737 organiza la Real Armada en tres escuadras "señalando para capitales de estas tres divisiones o departamentos los puertos de Cádiz, Ferrol y Cartagena"[8].

Los Departamentos de Marina se dividieron, a su vez, en nueve provincias o "partidos", cuya distribución inicial se recoge en la Tabla I. Finalmente, dentro de cada partido se establecía una distribución en poblaciones o municipios.

En cada Departamento se nombraba un "Intendente de Marina" de entre los oficiales de la Real Armada que, entre otras funciones, era la autoridad máxima de Marina y de la Matrícula. Éste, a su vez, nombraba "Ministros de Marina" en cada cabecera de partido, cuyas funciones eran la de garantizar los privilegios, establecer el Fuero de Marina, actuando como juez en las causas y, sobre todo, mantener actualizado el registro de las embarcaciones – pesqueras y comerciales - y de los matriculados. En este último caso las ordenanzas establecían un sistema exhaustivo de libros de registro donde se detallaban las circunstancias personales, los haberes recibidos y pendientes y el historial militar de cada individuo, clasificados según las distintas categorías de la marinería y de la maestranza.

El ministro de marina de cada partido nombraba dos aguaciles, para labores de policía, un letrado o auditor, que actuaría como abogado en las causas, y un escribano o notario exclusivo. Todos ellos gozaban del fuero de marina y, en su caso, tendrían la condición de oficiales de la Real Armada. A su vez, el Intendente nombraba en cada población un "subdelegado de marina" que estaba encargado de velar por todas las atribuciones de éste en cada municipio y que debía ser auxiliado por un abogado y un escribano público "de entre los existentes" que no estaban sujetos al Fuero de Marina, así como de un "cabo celador", de entre los matriculados, cuyas funciones eran las de llevar el control de los listados de matriculados, evitar conflictos entre éstos, velar por la defensa de sus privilegios – sobre todo

---

[5] Durante todo el siglo XVIII y principios del XIX, hasta su definitiva extinción en 1873, se desarrolló todo un cuerpo normativo de la institución, contenidas en las Ordenanzas de Marina de 1748, reformadas en 1751 y, posteriormente, en 1802, con las interrupciones impuestas por su derogación por las Cortes de Cádiz, renovada durante el Trienio Liberal de 1821-1824.

[6] Véase la Real Orden de 18 de enero de 1737, "*Ordenanza del Príncipe Almirante*", publicada el 21 de agosto de 1738.

[7] Véase el artículo LXXVIII del Título III del Tratado X de las "*Ordenanzas Generales de la Real Armada*" de 1 de enero de 1751.

[8] Cada división terminó recibiendo durante el siglo XVIII la denominación de "*Tercio*"; el "*Tercio de Levante*" o Departamento de Cartagena, el "*Tercio de Poniente*" o Departamento de Cádiz, y el "*Tercio del Norte*" o Departamento de Ferrol.

los relativos a la exclusión de los no matriculados de las actividades pesqueras y marítimas – y verificar el cumplimiento de la normativa vigente en materia pesquera[9].

Tabla 1. División de Partidos de cada Departamento de Marina en 1751[10]

El Ferrol	Cádiz	Cartagena
San Sebastián	Ayamonte	Vera
Bilbao	Sevilla	Cartagena
Santander	Sanlucar de Barrameda	Alicante
Ribadesella	Jerez de la Frontera	Valencia
Avilés	Cádiz	Tarragona
Vivero	Tarifa	Barcelona
Ferrol	Málaga	Mataró
Coruña	Motril	San Feliu de Guixols
Pontevedra	Almería	Palma de Mallorca

Fuente: Ordenanza de Matrículas de 1751, artículos I y II.

Pero, además, la Ordenanza de 1751 renovó o impuso, en su caso, un sistema de organización gremial en la pesca, restaurando estas instituciones de carácter medieval y/o adaptando las existentes a los fines requeridos. Cada gremio de armadores, patrones y marineros "sin mezcla de otros oficios que no gozan los Privilegios de Marina" debía elegir su propio cuerpo dirigente de "mayordomos, jurados, prohombres, clavarios o ministros" y organizarse para la defensa de sus privilegios mediante la elaboración de unos Estatutos, que debían ser visados por el Intendente de Marina del Departamento y establecer un sistema de financiación propio, que en general se sustentaba en el cobro de derechos sobre la pesca comercializada[11].

Libros de Registros de la Matrícula de Mar

Como hemos visto, una de las principales funciones de los "Intendentes de Marina" consistía en mantener actualizado el registro de todos los matriculados y, adicionalmente de todas las embarcaciones e instalaciones en las que éstos desarrollaban su trabajo habitual o en las que podrían estar incumpliéndose la exclusividad que éstos detentaban de las profesiones marítimas. De hecho. El artículo XXVI de la Ordenanza de 1751 establecía las siguientes funciones para los Ministros de Marina de cada provincia: "el gobierno, conocimiento, cuenta

[9] Algunos autores han considerado excesivo el control que estos "cabos celadores" impusieron sobre los matriculados, abusivo, en ocasiones e, incluso, partícipes de caso de corrupción y contrabando. Véase, por ejemplo, Vázquez Lijó, J. M., (2005).

[10] Esta división en partidos o provincias marítimas se mantendrá hasta la actualidad con cambios sustanciales. En el Tercio de Cádiz la provincia de Jerez es absorbida por Sanlucar muy pronto, en 1758, la capitalidad de la provincia de Tarifa pasa a Algeciras en 1786 y la de Ayamonte a Huelva por R.O. de 28 de mayo de 1835. En el Tercio de Cartagena la provincia de Vera es absorbida por la de Cartagena a principios del siglo XIX, de la provincia de Mallorca desgaja la de Mahón en 1786, que desde 1845 recibirá el nombre de provincia de Menorca, y también a principios del XIX las de Tortosa e Ibiza. La capitalidad de la provincia de San Feliu de Guixols pasa a Roses en 1786 y posteriormente a Palamós. Finalmente experimenta una profunda reorganización a principios del XIX, de forma que la provincia de Ribadesella es absorbida por la de Avilés cuya capitalidad pasa a Gijón. La provincia de Pontevedra se divide en dos nuevas provincias, la de Vigo y la de Villagarcía. Finalmente, en 1857, la capitalidad de la provincia de Vivero pasa a Ribadeo.

[11] Este control, limitado por los privilegios y franquezas otorgados a los matriculados, como veremos posteriormente, ha sido considerado, a nuestro entender de forma desproporcionada, como la causa principal de la práctica inexistencia de estadísticas pesqueras de capturas o producción en el siglo XVIII o de la reducida fiabilidad de las disponibles para el siglo XIX. Como veremos, los autores contemporáneos no opinaban de igual forma que los actuales.

y razón de la Gente de Mar matriculada; de las Maestranzas de Carpinteros de Ribera, y Calafates, de las Embarcaciones, que hubiere en las extensión de su Partido, la administración de justicia a todos éstos, y sobre negocio, y contratos marítimos; el cuidado del plantío, y conservación de los montes destinados a la cría de árboles de construcción, sus cortas, labras, y conducciones…" además de todas las competencias relativas a puertos, pesca, fletes, patentes de corso, naufragios, compras de materiales, etc. De hecho, desde el Reglamento de Comercio de 1723 se establece la obligatoriedad del registro de todas las embarcaciones dedicadas a actividades comerciales y pesqueras[12]. Dichos registros debían tener "noticia individual de todas las embarcaciones mayores y menores, con expresión de sus nombres, dueños, fábrica, medidas, etc, y empleos a que estén destinadas, llevando con ello una formal cuenta y razón en listas separadas, con distinción de puertos…" estableciéndose, además, el registro de todas las escrituras de trasmisión de embarcaciones[13].

Asimismo, sobre la Gente de Mar, los ministros debían llevar "listas exactas y claras de toda ella, con separación de lugares y distinción de hábiles y inhábiles"[14], dividiéndola en tres listas diferentes: la Gente de Mar, la Gente de Maestranza de Carpinteros y Calafates y la de jubilados que aún gozasen del fuero de marina pero no tuviesen ya que prestar servicio en la Real Armada. Estas listas o "libros maestros" debían estar foliados, con un folio destinado a cada marinero recogiendo expresamente sus datos personales: "…sus padres, lugar de nacimiento, edad, estado, y señales de rostro y cuerpo que lo hagan conocido…" y la información relativa a los servicios prestados: "…tiempo y paraje en el que se alistó en la matrícula, y sucesivamente con toda claridad sus destinos en los bágeles de Guerra o embarcaciones particulares, con expresión de clases en que haya servido y del modo en que conste se haya desempeñado, sus deserciones, castigos por delitos graves y generalmente, todas aquellas notas de consideración y dignas de ponerse en la lista…".[15]

Cada año el ministro de marina debía pasar una revista en todos los pueblos de su distrito, incluyendo en el informe "una nota o resumen del número de gente que resulte efectiva, con distinción de los que por las notas de sus asientos fuesen reputados hábiles, medianos o nuevos; el de los ausentes en conocido destino y el de los que, por ignorarse éste, hayan de considerarse desertores". Las listas de cada población se debían realizar siguiendo una misma estructura y metodología, de forma que los ministros debían velar para que todos los subdelegados "…observen igual método de claridad en sus listas, no permitiéndoles que en ellas pongan nota alguna hasta que el mismo les prevenga la que deban poner". Una copia literal de dichas listas e informes debía ser remitidos, al finalizar cada "revista", a la Contaduría principal del Departamento, que las debía disponer en tomos separados para cada población o distrito[16].

Las "inspecciones de matrículas"

La llevanza de los libros arriba mencionados no parecía ser suficiente. La Ordenanza de 1751 preveía que cada dos años se nombrase un "Ministro Inspector" en cada Departamento que, junto a un Capitán de Navío o Fragata, realizasen una revista de la gente matriculada, "… oigan y satisfagan sus quejas, pasen a la clase de inhábiles los que ya no estén de servicio,

[12] Véase Burgos Madroñero, M, (2003), pág. 24-25.
[13] Véase el artículo LXVI del Título III del Tratado X de las Ordenanzas Generales de la Real Armada.
[14] Ídem, artículo XXVII.
[15] Ibidem, artículo XXVIII.
[16] Desgraciadamente se han conservado muy pocos Libros de Registro de las Matrículas de Mar de las poblaciones del litoral, que podrían haber proporcionado una imagen realista de la evolución de la misma. Entre los existentes se encuentra los de la Matrícula de Mar de Santoña, depositados en el Archivo Histórico Provincial de Santander, según informan Coll y Fortea (1995), págs. 70 y 118.

propongan para goce de sueldo de invalidas los que le merezcan, reconozcan los puertos, montes, fábricas, etc., le informen de las utilidades que cada provincia produce o puede fácilmente producir en beneficio de la Marina, examinen la conducta de los Ministros y Subdelegados, les hagan oportunas prevenciones sobre la que convenga que observen en lo venidero y, al restituirse a la Capital del Departamento, enteren a su Intendente del verdadero estado de las dependencias de Marina en las provincias que hubieren visitado".

Pero a pesar de lo establecido en las Ordenanzas, la realidad parece que fue muy distinta, dado que para todo el periodo comprendido entre 1751 y 1802, año en el que se publica la nueva ordenanza de matrículas, sólo se realizaron algunas inspecciones en periodos muy concretos, de las que nos han llegado datos fragmentarios[17]. En la Tabla 2 se muestra la ubicación de los fondos archivísticos de las disponibles para el siglo XVIII y principios del XIX.

La estructura de la información proporcionada por las inspecciones fue variando con el tiempo, centrándose inicialmente en la mera comprobación del número de matriculados. Para cada Población y Provincia proporciona, en primer lugar, el número de matriculados hábiles clasificados en "Artilleros", "Marineros", "Grumetes", "Sirviendo en arsenales", "Sirviendo en buques de la armada o particulares", "Pilotos y prácticos" y "Cabos de matrícula", que totalizan los hábiles para el servicio. Seguidamente detalla los inhábiles o exentos: "Dueños de embarcaciones y patrones", "Jubilados" y "Dados por jubilados en la inspección". Asimismo se detallan los ausentes o no presentados: "Pasados a otro domicilio", "De ignorado paradero", "Excluídos de la matrícula", "Desterrados y presos", "Cautivos" y "Muertos". Respecto a los matriculados para la maestranza los clasifica en: "Oficiales de carpintero y calafates", "Obreros de carpintero y calafate" y "Aprendices". Para finalizar se detallan, puerto a puerto, las embarcaciones registradas existentes clasificadas según dedicación – pesca y comercio – y según tonelaje.

Tabla II. Inspecciones de Matrículas en el Siglo XVIII

Periodos	Situación de la documentación	Localización de la documentación
1752-1756	Desagregación completa	AGS, Leg. 300
1758-1765	Desagregación completa	AGS, Leg. 300
1772-1774	Sólo resumen de totales provinciales	AMDAB, Leg. 1873
1785-1787	Sólo resumen de totales provinciales	AMDAB, Leg. 1873
1795-1796	Agrupaciones de pueblos	AMDAB, Leg. 1883
1799	Agrupaciones de pueblos	AMDAB, Leg. 1883
1830	Desagregación completa	AMDAB, Leg. 1911
1832	Desagregación completa	AMDAB, Leg. 1914

Fuente: Elaboración propia.

[17] De hecho, según O'Dogherty (1952), sólo se realizaron para los Departamentos las correspondientes a 1765 y 1773. En un estudio de carácter local, Llovet (1980) localizó para la provincia marítima de Mataró las relativas a los años de 1754, 1765, 1774, 1786, 1795, y 1799, las mismas que menciona Burgos Madroñero (2003) analizando los fondos del Archivo Histórico de Marina para el Departamento de Cádiz, mientras que Fernández Díaz y Martínez Shaw (1984), analizando documentación del Archivo General de Simancas, proporcionan información de seis recuentos; los correspondientes al periodo 1752-1756, 1758-1765, 1772-1773, 1786, 1796 y 1799. Finalmente, en un trabajo más reciente, citando a Corroza (1866), a Salas (1870) y a Ocampo (1990), López Losa (2005) proporciona información correspondiente a las inspecciones de 1786-1787 y 1799.

Como hemos mencionado, los estados resultantes de las “Revistas de Inspección de Matrículas” no trataban de proporcionar una información exhaustiva del sector pesquero, pero de manera indirecta permiten inferir su importancia, analizar la tipología de artes utilizadas – según el tipo de embarcaciones existentes – y obtener una imagen fiel de la evolución de dicho sector a lo largo del siglo XVIII y, en ausencia de otra información, durante los primeros años del siglo XIX.

No obstante, conforme fueron manifestándose las causas que implicaban el mayor o menor desarrollo de la matrícula, se hizo evidente que para el crecimiento de ésta era imprescindible la expansión de las actividades pesqueras y, en menor medida, comerciales, dado que la capacidad de atracción de la matrícula, que garantizaba el privilegio del acceso exclusivo de los matriculados a estas actividades, estaba íntimamente relacionada con la capacidad de éstas para proporcionar un medio de vida digno a los potenciales tripulantes de la Real Armada.

Por ello, las inspecciones de finales del siglo XVIII centran su interés no sólo en verificar el número de matriculados, sino en describir las condiciones de vida de éstos, sus medios de subsistencia, las actividades económicas de las poblaciones y las industrias existentes, la situación del sector pesquero, los problemas y las posibles amenazas, así como realizando propuestas de mejora de la situación, incorporando, por tanto, en los informes abundante documentación acerca de la estructura de estas actividades.

Como ejemplo de ello pueden citarse los informes que Antonio Sañez Reguart, nombrado “Inspector General de Matrículas” y “Comisario Real de Guerra de Marina” en 1785, realizó durante la Revista de Inspección de Matriculas que éste dirigió entre 1785 y 1787. Sañez, a la sazón funcionario de Correos, se desplazó por todo el litoral, realizando informes detallados para cada población, inventariando las artes de pesca, tipos de embarcaciones y la problemática del sector, adquiriendo un conocimiento del mismo que le permitió publicar entre 1791 y 1795 lo que hasta el siglo XX ha sido la mejor y más detallada obra existente sobre la pesca en España: “El Diccionario Histórico de los Artes de la Pesca Nacional”[18]. En los informes que prepara de cada provincia marítima, que siguen una misma estructura, inserta propuestas de regulación, interviene en la resolución de conflictos planteados en el seno de los gremios por la competencia entre artes, realiza cuantificaciones de producciones, analiza la situación real de los pescadores y de las empresas de transformación y, en definitiva, da un paso más hacia las primeras estadísticas pesqueras. Como ejemplo de ello, en la correspondiente a la Provincia Marítima de Ayamonte, refiriéndose a La Higuerita, la actual Isla Cristina, detalla: “Se ocupan particularmente…por toda la temporada más de 2000 personas manufactureras. Giran anualmente dichas compañías en dinero efectivo que en embarcaciones propias y fletadas traen de Cádiz de 400 a 500000 pesos. Y en 1782 en que fue algo abundante la sardina excedió el consumo de la sal de 50000 fanegas”[19].

Los resultados del sistema

La evolución del número total de matriculados en el conjunto de los tres Departamentos de Marina, según las diferentes inspecciones, se recoge en la Figura 2. Desde la promulgación de

[18] Existe una edición facsímil de esta magna obra, que seguía el camino marcado en el país vecino por el “*Traité Géneral*” de Duhamel du Monceau (1769-1782) y por los artículos de “*L’Encyclopédie*” de Diderot y D’Alambert, realizada por el Ministerio de Agricultura, Pesca y Alimentación en 1988.

[19] Informe redactado el 4 de diciembre de 1787, *Provincia de Ayamonte, Séptima Revista del Departamento de Cádiz*, AMDAB, Legajo 1873.

las Ordenanzas de 1751 y hasta cerca de 1790 se mantiene un ritmo creciente en la incorporación de matriculados – aunque equívoco como veremos – alcanzándose un máximo en la inspección de 1785-1786. Posteriormente se produce una aguda crisis, de la que la matrícula no comienza a recuperarse hasta la década de 1820.

La razón es evidente; la expansión de la Matrícula de Mar entre 1750 y 1790 se produce en ausencia de conflictos bélicos marítimos importantes y en el contexto de una tremenda expansión de la Real Armada.

La evolución de la Real Armada durante el siglo XVIII y la primera mitad del XIX la recogemos en la Figura 1. Ésta pasó de los apenas 20 navíos y 17 embarcaciones menores existentes en 1751 a una flota integrada por 79 navíos, 57 fragatas y 178 embarcaciones menores en 1794, lo que implicaba – sólo de necesidades de marinería – requerir unos 8000 matriculados para el servicio permanente en la armada a unos 47000 tripulantes necesarios en 1794[20]. Frente a estas necesidades el número de matriculados había pasado, según las inspecciones, de 24.312 en 1755, a 35.493 en 1773 y 53.300 en 1786. Es decir, los matriculados cubrían la mitad de las tripulaciones necesarias para la expansión de la flota, pero dado que éstos, según las ordenanzas, sólo debían cumplir servicio un año de cada cuatro, la triste realidad no era otra que el número de matriculados existente proporcionaba, en su servicio ordinario, tripulantes para poco más del 25% de la flota. Pero, además, debe considerarse que dichas cifras de matriculados incluyen a los inhábiles (jubilados y enfermos), los exentos (patrones) y al personal de maestranza (carpinteros, calafates y toneleros), por lo que en realidad la situación fue mucho peor.

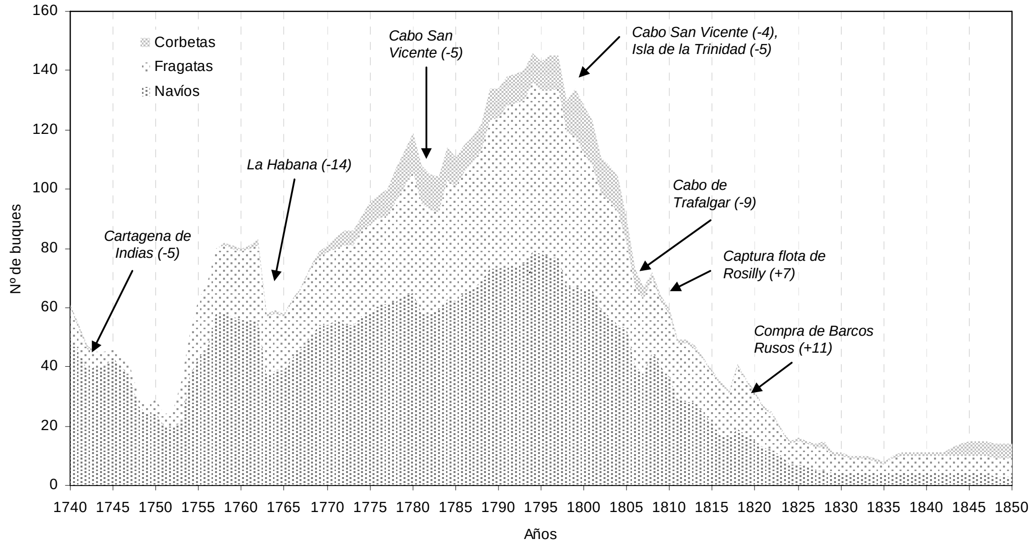


Figura 1. **Evolución de los buques de la Real Armada (1740-1850)**

En los periodos de especiales necesidades - cuando se armaban el máximo posible de navíos y fragatas - la Real Armada podía llamar a todos los matriculados y no sólo a la "cuadrilla" de dicho año, suponiendo una verdadera sangría en todos los municipios del

---

[20] Sálas, F.J. de, (1870), pág. 213, sobreestima las necesidades de marinería de la armada; asigna a cada navío el concurso de 800 tripulantes, cada fragata de 350 tripulantes, cada corbeta de 200 tripulantes y el resto de embarcaciones, en promedio, de 120 tripulantes. No obstante, según el Reglamento General de las Guarniciones y Tripulaciones de 1788, sólo los navíos de 112 cañones alcanzaban dicho nivel, siendo de 530 en los de 74 cañones – la mayoría – y de 472 tripulantes en los de 64 cañones. Una fragata de 36 cañones necesitaba 302 tripulantes y un bergantín 122. Pero no todos los tripulantes eran matriculados, dado que en estas cifras se integraban los oficiales de la armada y la guarnición de artillería e infantería, por lo que éstos eran, respectivamente, de 623, 360, 303, 214 y 122 matriculados.

litoral. El sistema podría haber funcionado en un contexto de conflictos bélicos cortos y esporádicos, pero en realidad los periodos de paz constituyeron la excepción: en los ochenta y siete años comprendidos entre 1739 y 1826 el Reino se encontró inmerso en situaciones de guerra contra Gran Bretaña, contra Francia o incluso contra ambas coaligadas durante sesenta años. Curiosamente, las inspecciones de matrículas que proporcionaron un mayor número de matriculados hábiles – la de 1752-1756, la de 1772-177 y sobre todo la de 1785-1787 – se realizaron durante los breves periodos de paz y arrojaron resultados satisfactorios. Pero a partir de 1793 y hasta 1816 se suceden los conflictos, exigiendo cada vez un mayor sacrificio a la Real Armada y provocando un fuerte descenso de los matriculados. En 1790, ante una amenaza de conflicto con Inglaterra, "…armó la España, en el corto tiempo de tres meses, al pie de sesenta y quatro navíos de línea, quarenta fragatas y cien buques de doce cañones hasta treinta: aún se hallaba capaz de armar diez o doce navíos y otras tantas fragatas; y todos estos buques se tripularon con marineros matriculados, de los que si todos no eran expertos, todos estaban acostumbrados a la mar…".[21] Para dotar dicha flota adecuadamente habrían sido necesarios un número total de matriculados cercano a los 35.000 marineros, que debieron permanecer en servicio por varios meses, durante los cuales, evidentemente no pudieron realizar las actividades propias en la pesca o el comercio, lo que unido a la falta de pago de los haberes, impuso un cambio drástico de actitud en los matriculados frente a sus obligaciones.

Como resultado cuando en "el año de 1793, se armaron nuevamente todas las fuerzas disponibles de la monarquía, y siendo algo menos numerosa que las que se movieron el año 90, ya faltó gente de mar para el total de sus tripulaciones, que hubieron de ser completadas con gente de leva"[22]. Pero el proceso continuó en los mismos términos. Tras la firma de la paz con Francia en 1795, se desarma la flota, y cuando en 1797 se inicia la guerra contra Gran Bretaña, tan sólo se consigue armar una flota de 27 navíos y 12 fragatas, y su comandante, el Teniente General José de Córdoba se lamentaba de que necesitaba "unos tres a cuatro mil hombres para tenerlos equipados con arreglo a las ordenanzas" para completar las tripulaciones; es decir, entre un 20% y un 25% por debajo de la tripulación reglamentaria. Finamente, en 1805, en vísperas de Trafalgar, Gravina indicaba que con los doce navíos que estaban armándose "nada se podrá realizar con estas fuerzas si no se presentan los cuatro mil y setenta y cinco individuos de marinería que nos faltan"[23].

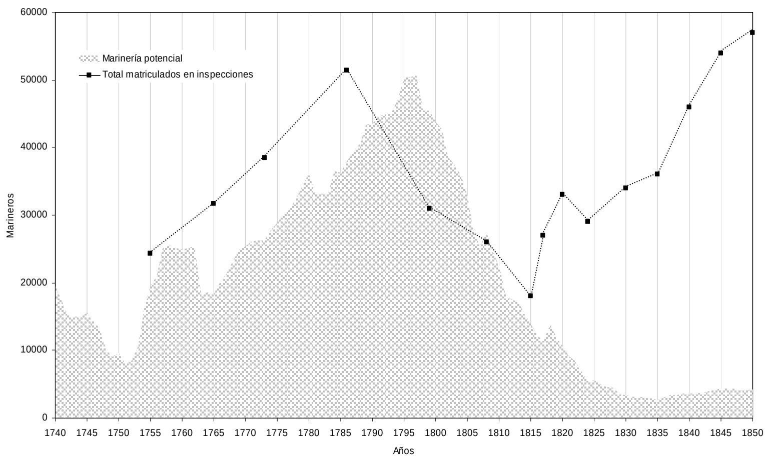


Figura 2.- Evolución del número total de matriculados según las Revistas de Inspección y de las necesidades de marinería en de la Real Armada

[21] Vázquez Figueroa, J., AMN, Mns. 1810-1811, Tomo II, citado en Salas (1870), pág. 219.
[22] Ídem, pág. 219.
[23] *Informe al Ministro de Marina* recogido en Lon Romeo, E., (2005), pág. 83.

Pero a pesar de la insuficiente matrícula y de los exiguos medios para mantener armada la flota, la política de construcción de nuevos barcos siguió manteniendo un ritmo frenético hasta finales del siglo XVIII, culminando en 1797 con la botadura de "El Argonauta", perdido en 1805 en Trafalgar, que sería el último gran navío de línea construido en España durante el siglo XVIII y la primera mitad del XIX[24]. Pero detengámonos un poco en analizar la situación a la luz que arrojan las inspecciones de matrícula. Considerando el Tercio Naval de Cádiz, entre 1755 y 1786 el número global de matriculados pasó de 10.401 marineros y personal de maestranza a 12.218, pero en el mismo periodo el número de desertores y dudosos se incrementó desde los 394 de 1755 a 2203 en 1799, mientras que el número de matriculados que se encontraban prestando servicio pasó de los 1132 de 1755 a los 4659 de 1799. En definitiva, la tasa de deserción evolucionó del 5.5% al 28%, mientras que de forma simultánea la tasa de matriculados hábiles en servicio, que se había mantenido entre el 11% y el 15% entre 1755 y 1786, se situaba en 1799 en el 60.7%, con las implicaciones que ello podía tener en las poblaciones del litoral[25].

Tabla III. **Evolución de la matrícula de mar en el Tercio de Cádiz**

	1755	**1765**	**1773**	**1786**	**1799**
Presentes o con licencia	5629	4476	5807	7485	814
En campaña	1132	920	1701	1339	4659
Desertores	394	2459	2142	2878	2203
Exceptuados	371	666	243	793	1072
Inhábiles	2875	2771	2147	2639	3470
Total matriculados	10401	11292	12040	15134	12218
Embarcaciones	1705	1743	1711	1899	2435

Fuente: Revistas de Inspección de Matrículas

Pero también resulta curioso analizar la evolución de las embarcaciones, que pasan de 1705 a 2435 en el periodo analizado, con un incremento relativo del 42.8% frente al 17.5% experimentado por los matriculados. El contraste es evidente, y se hace aún mayor comparando las dos últimas inspecciones, de forma que frente a un descenso del 20% de la matrícula, el número de barcos crece casi un 30%. Evidentemente estos resultados distan mucho de ser coherentes, y deben ser atribuibles a la deserción generalizada – de cada tres hábiles uno había desertado en 1799 – y a un fenómeno de carácter institucional. Entre los exceptuados de prestar servicio se encontraban los patrones de embarcaciones de más de 200 quintales, cuya evolución mantuvo un ritmo de crecimiento superior al de las

[24] De hecho, no se construyen nuevos navíos hasta 1852 cuando es botado el "*Isabel II*". De todos los astilleros de la armada entre 1797 y 1836 sólo se construyen seis fragatas y unidades menores. Como destaca Rodríguez González, A. R., (1999), pág. 8, tras la Guerra de la Independencia los astilleros estaban prácticamente abandonados: "*…el de Ferrol, por ejemplo, en el que trabajaban en 1790 uno 3500 obreros, en 1833 sólo tenía 37 hombres*".

[25] Los efectos de este estado de guerra permanente en los puertos pesqueros eran más que evidentes. La Real Cédula de 31 de marzo de 1805 "*considerando que con motivo de la presente guerra tendrán que salir de los puertos todos los matriculados útiles, y que quedarán sin ejercicio los mencionados barcos y aparejos, los pueblos sin pescado, las familias de la gente de mar sin arbitrios para subsistir*" tuvo que autorizar a los patrones a enrolar a terrestres en los barcos de pesca.

embarcaciones[26], de forma que si en 1755 eran matriculados el 21.8% de los propietarios, en 1799 dicho porcentaje se elevaba al 44%. Fernández y Shaw (2005) sugieren que "algunas embarcaciones se mantenían falsamente en uso para librar del sorteo a sus patrones"[27]. Evidentemente, ante el miedo a tener que realizar el servicio en pleno periodo bélico se utilizarían todos los medios, dado que no sólo era muy arriesgado el servicio en tiempos de guerra, sino que habitualmente el marinero se podía ver afectado por múltiples enfermedades y accidentes a causa de la permanencia durante periodos dilatados en el mar y a las penosas condiciones de vida a bordo[28]. Pero los inspectores, como hemos visto, también realizaban la revista de las embarcaciones, de las que existían exactos registros, por lo que dicho fenómeno debe ser minimizado y debemos atribuir el incremento de la flota mercante y pesquera a la favorable evolución económica de la segunda mitad del XVIII, atribuible en la primera por la liberalización de comercio trasatlántico desde 1768 y, en la segunda, por la expansión de la demanda debido al crecimiento de la población, de la reducción de las importaciones de bacalao inglés y la introducción de nuevas técnicas extractivas.

De esta forma, el número de matriculados mantuvo, en mayor o menor medida, el ritmo de la expansión de la flota hasta finales de la década de 1790 pero, posteriormente, se produce una aguda crisis manifestada en deserciones, desapariciones, ocultaciones y fraudes de forma que, cuando se llega al momento crítico de la Marina en vísperas de Trafalgar, ante la insuficiencia de matriculados, la mitad de las tripulaciones está compuesta por forzados y gentes de leva, "…hombres nada acostumbrados a la mar, y tan miserables, que ni aún tenían para evadirse de ella por medio del soborno, entraron a bordo de los navíos tan desnudos de ropa, como cargados de vicios.."[29]. La presión sobre los matriculados era ya insostenible, reduciéndose la producción pesquera como resultado de la sangría del sector. Como consecuencia, y a pesar del conflicto bélico con Gran Bretaña, las importaciones de bacalao casi se triplican en las dos últimas décadas del siglo XIX[30].

La causa de estos males fue muy evidente para los contemporáneos e incluso podría haber sido previsible. El que fuese Ministro de Marina en diversas ocasiones durante el primer tercio del siglo XIX, nos lo relata: "La marina real que formó el señor bailío Valdés me pareció siempre desproporcionada a la marinería que teníamos: era un gigante con una gran cabeza, piernas flacas y pies chicos y débiles. Y así, cuando en el año de 1790 y 91 nos presentamos en la mar en la gran parada con 50 a 60 navíos en diversos puntos, un correspondiente numero de fragatas y buques menores, pusimos, como suele decirse, toda la carne en el asador y, conforme fuimos experimentando contratiempos, no pudimos irnos reponiendo, no en lo material, sino en lo personal, y, si no hubiese cesado la causa política de aquel armamento y no se hubieren desarmado nuestras poderosas escuadras, ellas mismas por sí se hubieran desarmado por falta de individuos con que reemplazar los que iban faltando"[31].

En definitiva, debido a tremenda presión resultante de la amenaza que para la propia supervivencia del imperio colonial suponía el creciente poder marítimo de Gran Bretaña, el

[26] En la Inspección de 1773 hay varias provincias de las que no se proporciona el número de patrones o propietarios, como la de Tarifa o la de Almería, lo que explica el número excepcionalmente bajo contenido en la Tabla 3.
[27] Fernández Díaz, R. y C. Martínez Shaw, (1995), pág. 257.
[28] En Martín García (1999) se recoge un análisis de las partidas de defunción existentes en el *Archivo Eclesiástico de Marina* localizado en Madrid. De la muestra utilizada, con 1190 partidas correspondientes a tres expediciones concretas de 1776-1778, 1797-1803 y 1799-1801, tan sólo 8 fallecidos lo fueron en combate y 1033 por enfermedades.
[29] Vázquez Figueroa, J., (1811), T. II.
[30] Según Grafe, R., (2003), pág. 24, la entradas de bacalao en Bilbao, que se mantuvieron en torno a 60.000 quintales en el periodo prebélico anterior a 1790, llegan a alcanzar los 140.000 quintales a final del siglo.
[31] Véase Vázquez Figueroa, J., AMN, Mns. 432.

sistema cuya implantación recomendara un siglo antes el mercantilista Uztariz, había caído en un error anticipado ya por éste: "…pues serviría muy poco una Armada de muchos navíos, si no estuviesen gobernados con oficiales y demás individuos, capaces en la profesión, y hechos a los trabajos y peligros; siendo cierto que 20 bájeles con buenos Comandantes y tripulaciones, obrarán más que 40 con gente bisoña o de pocas experiencias…"[32].

Pero no había sido el único en anticipar el problema dado que el propio Marqués de la Ensenada informó al Rey Fernando VI en 1746 sobre la imposibilidad del país para disponer de una Armada comparable a la británica[33]. De hecho, diez años después de haberse manifestado con toda su crudeza el declive de la Real Armada y transcurridos casi cien años de las palabras del mercantilista, el Ministro de Marina se lamentaba ante las Cortes de Cádiz en estos términos "si hubiéramos tenido disponibles 8 navíos y 12 fragatas no más, es bien seguro que el servicio de tropas a América hubiera sido menos costoso y más rápido".

El declive de la Matrícula de Mar no comenzó a remitir hasta la finalización en 1828 de las crisis bélicas que se habían sucedido ininterrumpidamente desde 1793[34]. Finalizada la Guerra de la Independencia y las luchas coloniales, que provocaron la aparición del "corso insurgente" hasta en las mismas costas peninsulares, las inspecciones de matrícula muestran una clara recuperación, aunque los niveles de 1786 no se alcancen hasta 1845 como se muestra en la Figura 2[35].

El valor estadístico de las "Revistas de Inspección"

La cuestión más relevante acerca de la información contenida en las diferentes revistas de inspección de matrículas que se han conservado es si éstas pueden ser utilizadas para obtener una idea aproximada de la estructura del sector pesquero en España durante el siglo XVIII. Hay algunos trabajos que han explotado dicha información, como el ya citado de Fernández Díaz y Martínez Shaw (1984), y su conclusión principal es que aunque "los datos que más nos interesan son la evaluación de la gente de mar dedicada a esta actividad y el tonelaje de la flota pesquera…no tenemos relación directa de ninguno de estos parámetros"[36]. A pesar de ello, el método seguido por dicho trabajo, que analiza la revista de inspección de 1758-1765, parece acertado; estima a través de los puertos en los sólo hay flota pesquera el numero medio de tripulantes por embarcación, considerando los matriculados clasificados para marinería, y dicha media la aplica para obtener una estimación del número de pescadores en los puertos donde existe una mayor diversidad de actividades marítimas de un mismo Departamento. De esta forma obtienen para cada Puerto, Provincia y Departamento el número de pescadores y de embarcaciones dedicadas a la pesca, aunque admiten que la validez de dicha estimación "depende en cada región de los sistemas de pesca que en ella resulten predominantes, pues

---

[32] Uztariz, G., pág. 228.

[33] Ensenada manifestaba que "*proponer que V.M. tenga iguales fuerzas de tierra que la Francia y de mar que la Inglaterra seria delirio, porque ni la población de España lo permite, ni el Erario puede suplir tan formidables gastos*", y justificaba así una Armada menor – de 60 navíos y 65 fragatas - pero que junto a la Francesa pudiera hacer frente a la de Gran Bretaña [Citado en Salas, J.J., (1870), pág. 191.

[34] De hecho, aún en 1821 y 1822, tras la supresión temporal de la Matrícula durante el Trienio Liberal, en las convocatorias de matriculados realizadas por las Cortes para los tres Departamentos de Marina de los 3866 convocados tan sólo se presentaron 647 veinte meses después del llamamiento, reforzando la postura de aquellos que defendían su reinstauración [Salas (1870), pág. 274].

[35] En Gámez Duarte (2006) se desarrolla una extensa investigación acerca de la incidencia del corso de las repúblicas latinoamericanas sobre el comercio peninsular así como de los medios empleados para su frenar estas actividades.

[36] Fernández Díaz R. y C. Martínez Shaw, (1984), pág. 185.

cada uno de ellos emplea un determinado número de pescadores”[37]. Pero el principal inconveniente que debe plantearse sobre la representatividad de las estimaciones que puedan obtenerse a través de una técnica como la descrita no se encuentra relacionado tanto con la tipología de artes - que también - sino con el empleo de no matriculados en actividades pesqueras.

Hay dos excepciones de carácter legal a la exclusividad de la pesca para matriculados. La primera tendría poca importancia cuantitativa, pero mayor a la hora de interpretar una parte relevante de la información facilitada en las revistas de inspección, y reside en el hecho de que, como recoge el artículo LXV de las Ordenanzas de 1751, podrían existir propietarios de embarcaciones pesqueras o de tráfico que no fuesen matriculados, con la única condición de que no pudieran embarcarse. Pero la segunda excepción tiene una mayor relevancia. Tras reiterar en los artículos LXXXVIII y CXXI la exclusividad de la pesca a los matriculados, se introduce una matización bastante importante: ”…los pescadores matriculados podrán valerse a su arbitrio de gente no matriculada en todo lo que no pertenezca a la pesca, fuera de los barcos de ella, como en ayudar a tirar las redes a tierra, matar el pescado, salarle, etc. (cuando de la matriculada no haya la bastante para estos ejercicios, pues esta debe siempre emplearse con preferencia) entendiéndose la exclusión únicamente de navegar, como tales pescadores, en los barcos de pesca, y de pescar por si desde tierra con red o con otro instrumento que no sea vara o caña, cuyo genero de pesca a ninguno se prohíbe”. Y es aquí donde reside el principal inconveniente, dado que la mayoría de los artes de pesca existentes en la época eran artes playeros como las almadrabas de tiro, jábegas, sedales, boliches, lavadas y chinchorros, de los que con la excepción de los tres primeros – los de mayor tamaño – que requerían el concurso de una embarcación, los restantes se manejaban a pie en las playas y rías[38]. Y aún así, en el caso de almadrabas, jábegas y sedales, la normativa lo único que impedía era que los tripulantes de la embarcación que extendía el arte fuesen matriculados, pero no lo exigía para la mayor parte de los implicados en la faena; es decir, para los que ayudaban “a tirar las redes a tierra”.

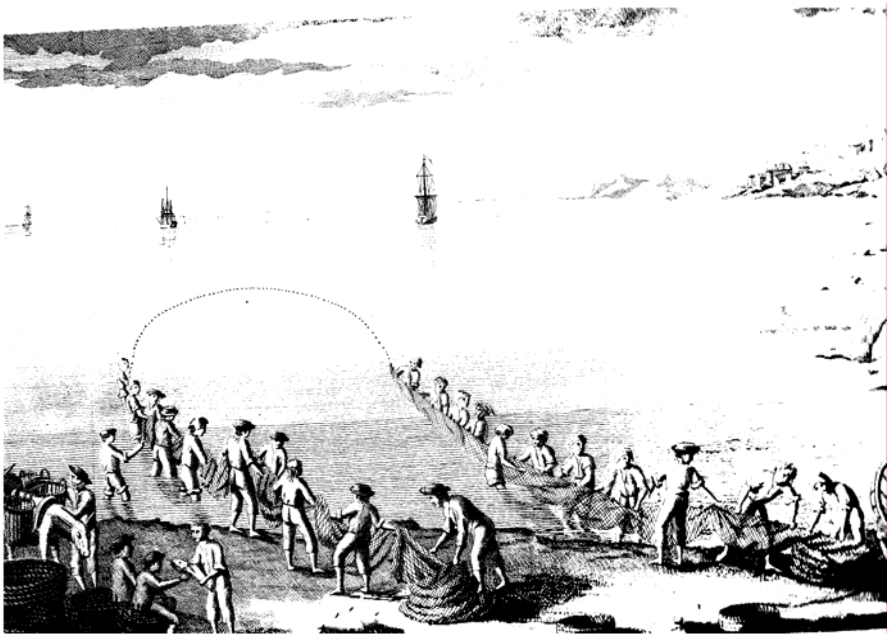


Figura 3. Pescadores sacando una jábega según Sañez Reguart (1791).

[37] Ídem, pág. 186.

[38] Para una descripción de estos artes y el número de trabajadores que empleaban puede consultarse García del Hoyo, J.J., (2002), pág. 26-38.

Siguiendo, por ejemplo a Sañez Reguart, en su descripción de la jábega se indica claramente la proporción de matriculados: "…el número de hombres de mar o matriculados que necesita una barca de xábega, es conforme las circunstancias locales o la costumbre establecida; a la que se añade la variación del mayor o menor tamaño de la barca y red. En unas partes se sirve un arte de estos con quince, con diez y ocho, y en otras con veinte o veinte y dos hombres. Estos son los que gobiernen la barca para el calamento de la red. Asimismo necesita otros tantos terrestres, a quienes se distingue con el nombre de gente de cabo de tierra, que son los que tiran de la xábega". Es decir, que en este arte, usualmente, existían tantos terrestres como matriculados y lo mismo sería acertado para el sedal. Los boliches no eran más que jábegas pequeñas, mientras que las lavadas y chinchorros, eran artes similares pero que se utilizaban preferentemente en rías y esteros - típicas de la costa de Huelva y Cádiz – y los botes o "baxeles" utilizados tenían sólo seis bancos con dos tripulantes – matriculados – en cada banco y un número proporcionado de terrestres en el cabo de tierra. Finalmente, en las almadrabas de tiro, dependiendo de su dimensión, podrían emplearse en promedio unos 300 trabajadores, de los cuales 75 tripulaban las 4 ó 5 embarcaciones requeridas para manejar el arte y serian matriculados mientras que el resto – terrestres – se dedicarían en su mayor parte al tiro de la red y a otras faenas relativas a la preparación y la venta de los atunes. Pero también existían otros artes en los que la práctica totalidad de los trabajadores debían ser matriculados. Los bous - iniciales artes de arrastre cuya introducción a mediados del XVIII fue una fuente constante de conflictos en todo el litoral español - los palangres, los aparejos de mano, y otras artes de red fijas como corvinales, trasmallos, sardinales, etc, eran actividades en los que se requería el estar embarcado para poder realizar la faena, por lo que estaría imposibilitado el acceso a estas modalidades de los no matriculados o "terrestres". De cualquier forma debe ser razonable utilizar un multiplicador cercano a dos por cada matriculado empleado en la pesca para estimar el número de trabajadores del sector, sin incluir a los requeridos para la preparación del pescado en las chancas. Pero disponemos de una pequeña memoria sobre la pesca de la sardina presentada por Juan Manuel de Oyarvide, en aquella época funcionario de la Real Hacienda en Huelva, a la "Real Sociedad Económica de Amigos del País de Sevilla" en 1776, en la que describe de forma detallada los rendimientos, las infraestructuras existentes, las embarcaciones requeridas y el empleo generado en la costa de Huelva durante dicha pesquería, diferenciando las embarcaciones foráneas de las locales y enumerando cada uno de los armadores y de los fomentadores asentados en la costa. En la temporada descrita, que sería la de 1775 dado que el informe se presenta en febrero de 1776, se utilizaron en la costa onubense unas 54 jábegas de las que 11 venían de la costa de Málaga, siendo el resto de las poblaciones de Ayamonte (29), Lepe (3) y Huelva (11), cifras prácticamente coincidentes con las que proporciona la Revista de Inspección de 1773. A éstas embarcaciones se unían 33 jabeques y 30 charangueros que actuaban como enviadas para transportar las capturas a las factorías de salazón o a los puntos de consumo, de los que 16 eran malagueños.

En conjunto, suponiendo que todos los tripulantes fuesen "gentes de mar", se requería el concurso de 1739 matriculados de los que unos 475 serían foráneos, más 1268 terrestres como "gente de cabo de tierra" y otros 821 terrestres en las labores de "espichado". Es decir, que la pesquería de sardina, que se extendía desde julio a enero, implicaba a 90 embarcaciones locales y a 1268 matriculados, prácticamente el 85% del total de matriculados disponibles. Debe tenerse en cuenta que la temporada de sardina representaba el periodo del año de mayor actividad pesquera, por lo que no resultan descabelladas dichas cifras. El multiplicador pescadores-matriculados resultante del análisis de Oyarvide, excluyendo a los trabajadores de las chancas de salazón, es del orden de 1.94 pescadores por cada tripulante matriculado, ratio que debería ser similar en todo el litoral de la provincia, dada la elevada presencia de artes playeros en la época.

Pero si en Huelva, donde las jábegas tan sólo representaban el 10.7% de la flota que puede identificarse como pesquera en la Revista de Inspección de 1773, el multiplicador es 1.94, resulta evidente que en Almería, donde éstas representaban el 84.4% de la flota pesquera, en Motril con un 78.4% o en Málaga con el 22.9%, dicho multiplicador debía ser aún mayor, empleando, por consiguiente, a un gran número de terrestres en las faenas de pesca. Por tanto, la cifra de matriculados, exceptuados los de maestranza, sólo puede servir como indicador de la dimensión del sector pesquero cuando sea corregida mediante un análisis pormenorizado de las actividad pesqueras que se desarrollaban en cada punto del litoral, y en ese sentido, debe ser tomada como un mínimo del empleo existente en el sector.

Las primeras estadísticas pesqueras

Como ya se ha comentado anteriormente, parecía evidente que la base fundamental para la propia existencia de una abundante matrícula de mar residía en el sector pesquero y, en menor medida, en el comercio marítimo. Por esa razón se habían establecido los derechos exclusivos de los matriculados y se velaba por su eficacia. Los informes de Sañez Reguart para la inspección de 1786 incorporan abundante material sobre la actividad pesquera, siendo en la correspondiente a la Revista de Inspección de 1795 cuando comienzan a sintetizarse dichas “noticias”, aunque aún de forma bastante rudimentaria. Pero a medida que las dificultades para el reclutamiento van siendo más acuciantes, el interés sobre el estado del sector pesquero – medio de vida de la inmensa mayoría de los matriculados – se convierte en una constante en las inspecciones y en objeto de los desvelos de los sucesivos secretarios y ministros de Marina, sobre todo tras el desastre de Trafalgar, que muchos habían atribuido al enorme porcentaje de “gente de leva” que había tenido que ser incorporada a las tripulaciones para que pudieran ser completadas. Paradójicamente, las primeras estadísticas exhaustivas surgen, precisamente, en el contexto más desfavorable, cuando el Estado prácticamente ha desaparecido, y se encuentra reducido a la ciudad de Cádiz, en plena Guerra de la Independencia. Tras la Batalla de Ocaña y la invasión de Andalucía, la Junta Central se refugia en Cádiz, donde se van concentrando funcionarios y militares que huyen de las tropas napoleónicas. Las presiones de las juntas provinciales, en especial de la sevillana, refugiada en Ayamonte, consiguen que la Central se disuelva tras la convocatoria de Cortes, haciéndose cargo del Gobierno un Consejo de Regencia, bajo la presidencia interina del general José Castaños y en el que participa Antonio Escaño. El Consejo designa un Gobierno en el que participa como funcionario del mayor nivel el oficial de la Real Armada José Vázquez Figueroa (1770-1855), ocupando interinamente la cartera de Hacienda y en propiedad la de Marina desde finales de 1810 a abril de 1813[39].

La situación en 1810 es desesperada. La autoridad real del Gobierno se limita a la propia ciudad del Cádiz y algún enclave costero como Tarifa y Ayamonte. Para la defensa de la plaza Vázquez se ve obligado a constituir una división de “fuerzas sutiles” y con los escasos recursos existentes, comenzar a armar algún que otro navío que pueda ser enviado a las

[39] Este oficial gaditano se había incorporado como guardiamarina en El Ferrol en 1788, donde fue profesor de Matemáticas, Cosmografía y Navegación. Como alférez de fragata participa en diferentes campañas desde 1793 obteniendo el grado de teniente de navío. En 1801 se incorpora al equipo de Domingo Pérez de Grandallana para elaborar las nuevas Ordenanzas Navales y las Ordenanzas de Matrícula de 1802. Por R.O. de 13 de noviembre de 1803 se le da de baja en el Cuerpo General de la Armada incorporándose como Oficial en la Secretaría de Estado y del Despacho Universal, situación en la que se encuentra aún en 1810 y donde acumuló una gran experiencia. Ocupó la cartera de Marina (1810-1812, 1812-1813 y 1816-1818) y la de Hacienda (1812). En 1818, a resultas del escándalo de la compra de las barcos rusos por la “Camarilla Real”, dimite y es exilado a Santiago. A la muerte de Fernando VII es llamado por la Regente para ocupar la cartera de Marina (1834-1835).

expediciones que se comienzan a mandar para sofocar las rebeliones iniciadas en las colonias americanas o colaborar en la lucha contra el ejército invasor.

Bajo su mandato, aunque no haya quedado constancia de la orden expresa, se encargan a los responsables de marina de las provincias bajo control del Gobierno la elaboración de una revista sobre el estado de las matrículas y de la pesca, dado que "el fomento de la pesca es uno de los medios más eficaces para acrecentar la marinería"[40]. El resultado son sendos estadillos remitidos por las provincias de Ayamonte y Algeciras, en los que se detalla el producto de la pesca y de las industrias de salazón, el capital invertido en éstas, el número y tipología de las embarcaciones pesqueras y de tráfico, la marinería existente y las retribuciones que disfrutan los inválidos de la matrícula[41]. En el caso de la provincia marítima de Algeciras se incorpora información sobre ésta y sobre Tarifa, Ceuta y San Roque, mientras que para la provincia de marina de Ayamonte, que corresponde a la actual provincia de Huelva, se facilita información de todos los distritos de marina de la misma; es decir, Huelva, que se extendía desde el Guadalquivir hasta la actual Punta Umbría, Cartaya, Lepe, La Higuerita y Ayamonte[42].

Estas son, sin lugar a dudas, las primeras estadísticas pesqueras españolas encargadas institucionalmente y dirigidas no a conocer el estado de la matrícula sino a analizar la situación del propio sector pesquero[43]. Responden a una misma iniciativa centralizada que surge del propio Vázquez Figueroa, dado que el mismo informaba en 1810 de la necesidad de "formar o establecer con otra parte de ella apostaderos al levante y al poniente (de Cádiz), en Huelva o Ayamonte, Tarifa o Algeciras, con el fin de sostener aquellos puntos, sacar víveres y convoyarlos a Cádiz"[44] y, posteriormente, en sesión de las Cortes de Cádiz de 24 de agosto de 1811, celebrada al efecto de suprimir las matrículas de mar, reivindica el mantenimiento de la institución y el reforzamiento del derecho exclusivo de pesca de éstos mediante la supresión de algunos tributos eclesiásticos que aún gravaban la pesca en Galicia, la prohibición de las artes de bou y la abolición del privilegio exclusivo de calar Almadrabas que la Casa de Medina Sidonia - en dicho momento el Marqués de Villafranca - disfrutaba desde la Edad Media, dado que debido a éste "quedan los marineros privados de la mejor y más abundante riqueza que produce una pesquería como la del atún, al mismo tiempo que ve engrosarse con ella a extraños en los tiempos precisamente más lucrativos del año, en los cuales la gente de mar ha de varar sus barcos y dejar pudrirse sus redes"[45]. Lógicamente, el que sólo se disponga de los datos correspondientes a estas provincias se debe a que eran las que, salvo breves incursiones francesas, se mantuvieron bajo control directo del Consejo de Regencia de forma permanente. Su importancia radica no sólo en que sean las primeras estadísticas pesqueras españolas, sino que su contenido servirá de modelo para las que se desarrollen durante más de un siglo.

---

[40] Vázquez Figueroa, J., AMN, Mns., 1810-1813, reproducido en Amorós, N., (1925), págs. 37-111.

[41] La denominación de los mismos es elocuente: "*Estado que manifiesta el número de embarcaciones de todos portes y tráfico, las clases de pescas y sus productos, fábricas de pescado y sus clases, capitales que giran los traficantes de este ramo y el número de marineros acreedores de premios y los que disfrutan de inválidos correspondientes a la provincia de Ayamonte*" (AMDAB, Legajo 1984).

[42] La capitalidad de la provincia marítima de Tarifa había pasado a Algeciras en 1787, mientras que Ayamonte pierde la capitalidad a favor de Huelva por R.O. de 28 de mayo de 1835.

[43] Los originales de los documentos mencionados se encuentran recogidos en el Legajo 1984 del Archivo Histórico de Marina "Alvaro de Bazán" y se recogen en por Burgos y Lacomba (1993) y Burgos (1992).

[44] Véase Amorós (1925), pág. 47.

[45] Véase el "*Discurso sobre las matrículas de mar pronunciado ante las Cortes generales y extraordinarias del Reino en sesión pública de 24 de agosto de 1811*" [Salas (1870), págs. 263-271].

La consolidación de unas estadísticas pesqueras

Las vicisitudes del reinado de Fernando VII, con la pugna entre el absolutismo y los liberales, la bancarrota derivada de los conflictos bélicos y la deficiente gestión de los sucesivos gobiernos impidieron la normalización del país. Prácticamente, hasta la definitiva asunción de la pérdida de las colonias americanas, la principal finalidad de los sucesivos gobiernos no fue otra que buscar los medios de enviar expediciones militares para reforzar los enclaves realistas que aún quedaban en ultramar. Con la caída de Chiloé y El Callao en 1826, con la excepción de la fallida expedición a México del brigadier Barrada en 1829, prácticamente se dan por terminadas las guerras coloniales. En dicho momento ocupaba la cartera de Marina Luis María Salazar Salazar, cuya trayectoria profesional y política, se entremezcla con la José Vázquez Figueroa, quién le introdujo en política, a quién sucedió en el cargo, por quién sería sucedido, y como aquel, también fue exilado tras su cese en 1816 por Fernando VII. A pesar de su adscripción inicial al ala moderada del partido liberal, fue progresivamente derivando hacia posturas más conservadoras terminando en el absolutismo más radical, siendo, de hecho, el Ministro que mayor permanencia tuvo en el Gobierno tras la restauración absolutista en 1823, manteniéndose en éste hasta poco antes del fallecimiento del monarca, cuando la Reina María Cristina se hace cargo de la Regencia y destituye al último gobierno absolutista.[46]. Independientemente de los matices políticos, la larga gestión de Salazar al frente del Ministerio de Marina proporcionó algunos frutos. De hecho, renovó parcialmente los arsenales y se botaron las primeras grandes unidades – fragatas - desde 1801[47]. Analizó intensamente la problemática de las matrículas, restaurándolas tras su supresión a comienzos del Trienio Liberal[48], e impulsándolas, de forma que durante su mandato el número de matriculados creció un 25%, situándose en 55036 marineros – resultando hábiles para el servicio 35995 marineros – y 6209 de maestranza. En este contexto surgen en España, las primeras estadísticas pesqueras, sistematizadas y en serie regular.

Descripción y contenido de las Estadísticas de los Productos de la Pesca

Bajo la denominación de "Estado General que demuestra los productos de las pesquerías del Reino en el año de 1831, formado por la Dirección General de la Real Armada con los datos reunidos al Efecto" se publicó en el Estado General de la Armada de 1833 un "estado de los productos de la pesca anual, beneficiada en las costas de España e Islas adyacentes, que por primera vez se publica, carece aún de toda regularidad que el tiempo y la constancia de las disposiciones podrá darle". Evidentemente era el primer "Estado General" que veía la luz

[46] Al igual que Vázquez Figueroa, Salazar (1758-1838) procedía del Cuerpo de Oficiales de la Armada, sentando plaza de guardia marina en Cádiz. Participó en diferentes campañas entre 1775 y 1793, siendo destinado como Oficial, igualmente, en la Secretaría de Estado y Despacho de Marina, participando, asimismo, en la redacción de las Ordenanzas de Matrículas de 1802, año en el que es ascendido a Capitán de Navío y destinado, nuevamente, al servicio en la Real Armada. En 1808 es nombrado inspector de matrículas, puesto que no pudo desempeñar a causa de la Guerra, instalándose durante ésta en Córdoba, Cádiz y, finalmente, en Galicia, de donde es requerido en 1812 por Vázquez Figueroa para que ocupase la cartera de Hacienda, iniciando su carrera política que le llevará al Ministerio de Marina (1814-1816, 1820, 1823-1832) y al de Hacienda (1814) además de haber ocupado otros interinamente.

[47] Desde la botadura del navío Argonauta en 1798 y de la fragata Prueba en 1801, tan sólo se construye en España el Bergantín "El Ferrol" en 1820 hasta la construcción de las fragatas "Iberia" y "Lealtad" en 1825 y la "Villa de Bilbao"en 1826. No se construye ningún navío hasta la botadura en 1851 del "Isabel II" y del "Francisco de Asís".

[48] La supresión se realizó mediante R.D. de 28 de octubre de 1820, cediendo el reclutamiento para la Armada a los Ayuntamientos, lo que constituyó un verdadero fracaso, como se ha indicado en la nota 33.

pública, pero no fue el primero que se elaboraba ni sería el último[49]. El "Estado" era el resultado de agregar los datos correspondientes al último semestre del año anterior y el primero del año en curso, considerados los semestres de diciembre a mayo y de junio a noviembre, sistema que se mantuvo hasta finales del siglo XIX, aunque en algunas provincias marítimas – como en Málaga - se realizaban estadillos trimestrales que, después, se agregaban en términos anuales, aunque ello planteaba algunos problemas que posteriormente comentaremos.

La información contenida era recopilada por las Ayudantías de Marina de cada Distrito partiendo de los datos facilitados por los gremios de matriculados, quienes las elaboraban utilizando los registros existentes del cobro de los arbitrios gremiales - "un tanto proporcional de la pesca que los pescadores benefician" - con los que se financiaba la corporación. Los estadillos de cada provincia marítima eran remitidos a la Dirección General de la Real Armada que se encargaba de realizar el "Estado" en el que se proporcionaba información agregada espacialmente para cada Provincia Marítima y cada Departamento de Marina. Desde el principio, la pretensión de la Dirección General de la Real Armada era elaborar estadísticas anuales que permitiesen comparar la evolución interanual, como se informaba en el propio documento[50].

Tabla IV. **Contenido de los Estados Generales de Pesca**

Variable	Desde	Unidades
Peso del Pescado cogido	1829-1892	Quintales (1829), Arrobas (1831-67), Kg. (1868-92)
Valor del Pescado Cogido	1829-1892	Rs. vn. (1829-63), Escudos (1863-68), Ptas. (1872-92)
Peso del Pescado Salado	1829-1892	Quintales (1829), Arrobas (1831-67), Kg. (1868-92)
Valor del Pescado Salado	1829-1892	Rs. vn. (1829-63), Escudos (1863-68), Ptas. (1872-92)
Peso del Pescado Escabechado	1829-1892	Quintales (1829), Arrobas (1831-67), Kg. (1868-92)
Valor del Pescado Escabechado	1829-1892	Rs. vn. (1829-63), Escudos (1863-68), Ptas. (1872-92)
Peso del autoconsumo de pescado	1829-1892	Quintales (1829), Arrobas (1831-67), Kg. (1868-92)
Peso del consumo local de pescado	1829-1892	Quintales (1829), Arrobas (1831/67), Kg. (1868/92)
Peso de las Exportaciones de salazones	1829-1892	Quintales (1829), Arrobas (1831/67), Kg. (1868/92)
Peso de las Exportaciones de escabeches	1829-1892	Quintales (1829), Arrobas (1831/67), Kg. (1868/92)
Peso de las ventas nacionales de salazones	1829-1892	Quintales (1829), Arrobas (1831/67), Kg. (1868/92)
Peso de las ventas nacionales de escabeches	1829-1892	Quintales (1829), Arrobas (1831/67), Kg. (1868/92)
Embarcaciones empleadas	1829-1892	Nº total de embarcaciones
Matriculados empleados	1829-1892	Nº total de matriculados (1829/72) pescadores (1872/92)
Fanegas de sal consumidas	1858-1892	Fanegas (1829/74), Kg. (1883/92)
Valor de las embarcaciones	1858-1892	Rs. vn. (1829/62), Escudos (1863/68), Ptas. (1872/92)
Valor de las artes empleadas	1860-1892	Rs. vn. (1829/62), Escudos (1863/68), Ptas. (1872/92)

Fuente: Elaboración propia.

La información facilitada tenía la estructura contenida en la Tabla IV, y adolecía, evidentemente, de graves inconvenientes, dado que no desagregaba por especies ni por tipo de arte o tipo de embarcación. No obstante, como veremos, dicha información era comparable a la que en la época se confeccionaba en otros países de nuestro entorno más inmediato.

[49] Este documento, por su importancia histórica, fue reproducido en Salas, J. y F. García Solá, (1876) en anexo sin paginar.
[50] A pié de tabla se afirma que "*presentará en adelante los productos, si no efectivos, al menos proporcionales, de los años comparados*".

La serie disponible del "Estado General de la pesca"

Pero el "Estado" publicado en 1831 no fue el primero confeccionado, sino que durante algunos años anteriores se habían realizado trabajos análogos que no habían sido publicados. El primero se realizó con los datos del segundo semestre de 1828 y primer semestre de 1829, facilitando la información sobre pesos en quintales y fue terminado el 27 de marzo de 1830, y su título era "Estado General demostrativo de la Pesca Beneficiada en las Costas de los Tercios de Poniente, Levante y Norte, con noticia de sus valores y cantidades saladas, escabechadas, exportadas e internadas al Extranjero y nuestro Reino en el año de 1829"[51].

Con posterioridad al publicado se confeccionaron los "Estados" correspondientes a 1832, 1834 y 1835, faltando, al menos en la documentación analizada, el correspondiente a 1833[52]. No obstante parte de estos "estados" están incompletos, con lagunas para el primer semestre de1835, además de la información de varias provincias marítimas[53].

En años siguientes se publicaron varios resúmenes anuales en la publicación oficial de carácter anual del Ministerio de Marina "Estado General de la Armada", como se recoge en la Tabla V, incorporando, de forma progresiva, información adicional, variando las unidades de peso – en quintales, arrobas y kilogramos - y monetarias – reales de vellón, escudos de vellón y pesetas – como se muestra. En esta publicación se difundieron los comprendidos entre 1831-1857.

Al iniciarse la llamada "Estadística Oficial Española" y comenzar a publicarse el "Anuario Estadístico de España" se incorpora en éstos la publicación de los "Estados de Pesca", y dejan de ser difundidos en el "Estado General de la Armada". En los cinco anuarios originales, correspondientes al periodo 1858-1867, se publica información de diez estados de pesca aunque, desgraciadamente, sólo para los cuatro primeros (1858-1861) se publicó la información desagregada por provincias marítimas, mientras que para los restantes se limitó a los datos globales por Departamento o Tercio de Marina. Lamentablemente, no se publicaron más "Anuarios Estadísticos" hasta principios del siglo XX., con la excepción de la "Reseña Geográfica y Estadística" de 1888.

En 1868 se constituye la Comisión Permanente de Pesca, como órgano de asesoramiento y participación del sector en el diseño de políticas y regulaciones. Dicha institución publicó sendos "Anuarios" en los que se difundieron "Estados" de algunos años y distritos concretos. La información era, en éste momento, muy deficiente y con importantes lagunas. Debe considerarse que desde 1867 se extinguen los gremios de mar, por lo que se liberaliza, parcialmente, el acceso al sector pesquero. En las estadísticas se percibe claramente el efecto de este cambio institucional – el número de pescadores crece en un 76% mientras que la flota crece un 50% entre 1867 y1878[54].

---

[51] Dicho "Estado" se encuentra depositado en el legajo 2128 del Archivo Histórico de Marina "Álvaro de Bazán" (AMDAB), y en el mismo no aparece información acerca de las provincias marítimas vascas ni de Canarias, lo que será frecuente en años sucesivos. En el mismo legajo se encuentran los cuadros auxiliares y el resumen del "Estado" de 1831 publicado.

[52] En las "*Memorias*" de Vázquez Figueroa, depositadas en el Archivo del Museo Naval de Madrid, se contiene una referencia a un informe remitido al Ministro de Hacienda sobre el "*Estado General de la Pesca*" de 1834, comparando sus resultados con el del año anterior (AMN 0247, Mns. 0455/063), por lo que sí se elaboró la estadística para dicho año.

[53] Dicha información se encuentra en el legajo 2129 del Archivo Histórico de Marina. El primer trabajo en el que se cita el contenido de dichos legajos es el de Burgos y Lacomba (1992).

[54] Los Anuarios de la Comisión Permanente así como las Memorias de Industria y Legislación, contenían información muy diversa; aparte de algunas estadísticas, incorporaban informes, memoriales, noticias sobre la pesca en otros países, recopilaciones de normativas y reseñas de las actuaciones administrativas.

Tras la restauración borbónica la Comisión Permanente de Pesca es sustituida por la Comisión Central de Pesca que difundía sus trabajos en la publicación "Memorias de Industria y Legislación de Pesca". Se publicaron tres memorias distintas; la del periodo 1870-1874, la correspondiente a 1874-1879 y, finalmente, la dedicada al periodo 1880-1884[55]. En estos trabajos se difunden algunos "estados de pesca" parciales, y en el segundo de ellos se incorpora el resultado de una investigación diferente realizada en 1878 un "estadillo", en el que se realiza un inventario de los medios de producción de todo el litoral, tanto propiamente pesqueros como de la industria derivada, pero no proporciona información sobre volúmenes de producción. En la "Reseña Geográfica y Estadística de España", publicada en 1888, se insertó el "Estado" correspondiente a 1883, que volvía a recuperar el formato tradicional. Finalmente, dichas publicaciones son sustituidas por la "Revista de Pesca Marítima", que se publicó de forma continua entre 1885 y 1901, en la que se insertaron dos "Estados de pesca" correspondientes a 1889 y 1892, con lo que finaliza la serie. Éste último se publicó en 1895 y, como novedad, además del resumen anual, facilitaba el detalle trimestral de cada distrito.

Tabla V. "**Estados Generales de la Pesca" localizados**

Periodo	**Localización**	**Situación**
1828-1829	AMDAB Leg. 2128	Completo Distritos
1830-1831	AMDAB Leg. 2128, EGA1833	Completo Distritos
1831-1832	AMDAB Leg. 2129	Completo Distritos
1832-1833	AHMN, MMSS 447	Referencias totales
1833-1834	AMDAB Leg. 2129	Completo Distritos
1834-1835	AMDAB Leg. 2129	Parcial Distritos, semestres
1845-1846	AMDAB, Leg.2131, EGA 1847	Completo Distritos
1846-1847	AMDAB Leg. 2131, EGA 1848	Completo Distritos
1848-1849	EGA 1850	Completo Distritos
1849-1850	Referencia en López Losa (2000)	
1850-1851	EGA 1852	Completo Distritos
1851-1852	EGA 1853	Completo Distritos
1852-1853	EGA 1854	Completo Distritos
1856-1857	EGA 1858, AEE 1858	Completo Distritos
1857-1858	AEE 1859-1860	Completo Distritos
1859-1860	AEE 1860-1861	Completo Distritos
1860-1861	AEE 1860-1861	Completo Distritos
1861-1862	AEE 1862-1865	Completo Departamentos
1862-1863	AEE 1862-1865	Completo Departamentos
1863-1864	AEE 1862-1865	Completo Departamentos
1864-1865	AEE 1866-1867	Completo Departamentos
1865-1866	AEE 1866-1867	Completo Departamentos
1866-1867	AEE 1866-1867, ACPP 1868	Completo Departamentos
1867-1868	ACPP 1869	Incompleto Distritos
1872-1873	MILP 1870-1874	Incompleto Distritos
1873-1874	MILP 1870-1874	Incompleto Distritos
1883	RGEE 1888	Completo Distritos
1889	RPM 1890	Completo Distritos
1892	RPM 1894	Completo Distritos

AMDAB:	Archivo Histórico de Marina "Álvaro de Bazán"	MILP:	Memoria de Industria y Legislación de Pesca
EGA:	Estado General de la Armada	RGEE:	Reseña Geográfica y Estadística de España
AEE:	Anuario Estadístico de España	RPM:	Revista de Pesca Marítima
ACPP:	Anuario de la Comisión Permanente de Pesca	AMN:	Archivo del Museo Naval

[55] Debieron imprimirse muy pocos ejemplares de estas memorias, dado que los existentes son muy escasos. De la última tan sólo existe un único ejemplar depositado en la Biblioteca del Senado.

Las estadísticas de pesca posteriores, que se inician en 1904, tienen ya una estructura radicalmente distinta, mucho más desagregada espacial, temporal y funcionalmente y, evidentemente, mucho más dirigida hacia la creciente demanda de información fidedigna que el incipiente "Instituto Español de Oceanografía" requería[56].

En definitiva, considerando la información publicada o no publicada que ha sido localizada, ya sea parcial o completa, se realizaron, al menos, veintiocho "Estados de Pesca" para el periodo 1829-1892, con la misma estructura definida durante el mandato ministerial de Salazar[57], aunque es probable que existan al menos otros dos: el correspondiente a 1846-1847, citado por López Losa (2000), o el de 1881, citado por Fernández y Navarrete (1905).

La primera laguna en la serie se produce, como hemos mencionado, en el correspondiente a 1829-1830, que no consta en el Archivo de Marina y debería haberse confeccionado entre diciembre de 1830 y enero de 1831. Pensamos que, en realidad, se confeccionó, pero aún no nos ha llegado ninguna información sobre el mismo. Posteriormente, no se ha localizado ningún otro "Estado" elaborado entre el de 1834-1835 y el de 1845-1846, es decir, un total de diez años faltan en la serie, pero corresponden, curiosamente, al periodo en el que el Ministerio de Marina es sustituido por el Ministerio de Marina, Comercio y Gobernación de Ultramar, por lo que es posible que los fondos relativos a la elaboración de las estadísticas de dicho periodo no fuesen transferidos a Marina, una vez separado el Ministerio en 1847, y que lo hayan sido al nuevo Ministerio de Comercio, Instrucción y Obras públicas, entre cuyos fondos podría encontrarse nueva documentación al respecto que aún pudiera ser localizada[58].

La serie contiene después algunas lagunas esporádicas, las correspondientes 1847-1848 y a 1849-1850, aunque López Losa (2000) parece haber utilizado datos de este último "Estado". No hemos encontrado, asimismo, referencia alguna del periodo 1853-1856, y no hay razón aparente para que el Ministerio de Marina no hubiese seguido elaborando la estadística, dado que los anteriores y posteriores a los periodos citados siguen publicándose en el "Estado General de la Armada". Diferente puede ser la causa de la existencia de lagunas en la serie para el periodo posterior a 1856. Por Real Decreto de 3 de noviembre de dicho año se crea la "Comisión de Estadística"[59], dependiente de la Presidencia del Consejo de Ministros, fecha que para muchos significa el comienzo de la "Estadística Oficial" en España[60]. Pocos días después, el 28 de noviembre, se publicó el Reglamento de la Comisión, en cuyo artículo 20 se

---

56 Estas se desarrollan a partir de la R.O. de 25 de junio de 1904 (BOE nº 74, p. 754).

57 Es posible que futuras investigaciones permitan localizar otros "Estados de Pesca", ya sean globales o parciales, en distintos archivos. En algunas poblaciones parte de los fondos de los archivos de las Capitanías Marítimas se han incorporado a los archivos municipales o a los Archivos Históricos Provinciales, donde puede ser que se localicen los datos originales remitidos en su día. También es probable en el Archivo General de Simancas o en el Archivo "Álvaro de Bazán" se localicen nuevos estadillos. Dado que tenemos constancia de que se remitieron copias de estos "Estados" al Ministerio de Hacienda, en su archivo, ya sea en el fondo histórico aún depositado en el Ministerio o en los remitidos al "Archivo Histórico Nacional" se pueda encontrar nueva información que permita ampliar la serie.

58 En las respuestas al "*Interrogatorio remitido por el Ministerio de Marina sobre el fomento de la industria de la Pesca*", promovido por el oficial mayor del Ministerio de Marina Jorge Lasso de la Vega se encuentra abundante material para deducir que la información continuó siendo remitida al Ministerio en este periodo. Por ejemplo, en la respuesta que realiza el Comandante de Marina del Distrito de Sanlúcar de Barrameda, se recoge información sobre la pesca capturada y su valor entre 1831 y 1846, coincidiendo los años correspondientes a la incluida en los "*Estados*" disponibles, y mencionando que "*la noticia que antecede sacada de los Estados semestrales de la pesca desde principios de 1831 hasta fin de 1846 cuyas copias fueron remitidas a la superioridad*" AMN, Mns. 2203, fol. 81.

59 Por Real Decreto de 15 de julio de 1865 la Comisión se transforma en la Junta General de Estadística y se crea una Dirección General de Estadística dependiente de Presidencia.

60 De hecho Sanz Serrano (1956), pág. 138, sugiere que el, por aquel entonces, Presidente de Gobierno, Ramón María Narváez "*puede considerarse como el verdadero fundador de la Estadística Oficial en España*".

establecía que correspondía "a la Segunda Sección", entre otras materias, "las fuerzas militares de mar y tierra". Lógicamente, la primera publicación donde se incorpora información sobre las matrículas de mar, en general, y sobre la pesca, en particular, fue el "Anuario" de 1858, donde se publica el "Estado de la Pesca" correspondiente al año 1857-1858. Parece, por tanto, que el hecho de que la Comisión fuese a encargarse de publicar dicha información fue la causa de la interrupción de su publicación en la serie "Estados Generales de la Armada". La fuente citada en los anuarios sigue siendo el Ministerio de Marina, que continuaría elaborando las estadísticas, pero no su publicación, dado que la difusión de los mismos queda atribuida a la Comisión. El último "Anuario" de esta primera serie fue publicado en 1870, correspondiendo a 1866-1867, donde se incorporó la serie de resúmenes de los estados de pesca desde 1861 a 1867.

Como ya hemos comentado, no hemos localizado ningún"Estado" completo para el periodo que se extiende desde lo publicado en este último Anuario y la publicación de la "Reseña Geográfica y Estadística" en 1888. Y ello no resulta casual por varias razones. En primer lugar por los avatares de la revolución de 1868 y, sobre todo por reformas administrativas relevantes: la conversión en órgano meramente asesor de la Junta General de Estadística, la adscripción de la Dirección General de Estadística al Ministerio de Fomento el 26 de abril de 1870 y, finalmente, la supresión de ésta por Real Decreto del 4 de agosto de 1871 incorporándola como sección o negociado a la Dirección General de Agricultura, Industria y Comercio, donde "la Estadística no era más que una pequeña parte de los asuntos en que debiera intervenir"[61].

Pero también hay razones de carácter institucional. En primer lugar, por R.D. de 10 de julio de 1864 los gremios de mar se declaran extinguidos, comenzando a resquebrajarse, lentamente, la estructura sobre la que había reposado la elaboración de las estadísticas[62]. Las Comandancias no pueden acudir ya a las casetas o tinglados que en las playas tenían los gremios para el pesado y venta del pescado, que con sus limitaciones y probables tasas de ocultación, "se aproximaban a lo cierto más que los que en la actualidad se emplean", que no eran otros que declaraciones de los patrones y armadores de los artes[63]. Y, además, tras diversas vicisitudes, mediante R. D. de 22 de diciembre 1873 (G.M. de 17 de enero de 1873), durante el breve reinado de Amadeo de Saboya, se autorizó a las cortes a tramitar un proyecto de ley aboliendo las matrículas de mar, lo cual tuvo efecto, una vez proclamada la Primera República mediante Ley aprobada por las Cortes el 2 de marzo de 1873 (G.M. de 26 de marzo de 1873), estableciéndose entonces que "el ejercicio de las industrias marítimas es libre para todos los españoles" y entendiéndose por industria marítima "el tráfico de puertos y la pesca en general". La institución creada en 1605, y reforzada a partir de la Ordenanza del Príncipe Almirante en 1737 había finalizado su existencia. Sin gremios, sin matriculados, y en plena crisis política, lo extraño es que aún existiesen"Estados de Pesca" parciales para el periodo 1873-1874, lo que debe ser atribuido a la mera inercia del aparato administrativo de Marina[64]. Posteriormente parece que estas estadísticas, con la excepción de las correspondientes a 1883,

[61] Idem, pág. 165.

[62] Tras la supresión inicial de los gremios por las Cortes de Cádiz, la mayoria, mermadas sus competencias, perviven hasta esta época. Como reconocía García Solá (1880), pág. 834, "*desde que la supresión de los gremios encargados de suministrar estos datos la dejó sin medios de poderla continuar, inútiles han sido las disposiciones dictadas excitando el celo de las autoridades locale*s".

[63] Ante esta situación la Comisión Permanente de Pesca tuvo que imponer a los Cabos de Matrícula la obligación de "*llevar una cuenta exacta del pescado que diariamente desembarca en los muelles*", refiriéndose a la provincia marítima de Cartagena. Véase Fernández Duro, C., (1868), pág. 206.

[64] De hecho, cuando Mariano Carreras y José Piernas Hurtado editan en 1873 su "*Tratado de Estadística*" el último dato que proporcionan (págs. 281-282) corresponde a 1867.

1889 y 1892, no se realizan o, al menos, si se realizaron, muy pocos tuvieron acceso a las mismas[65].

De cualquier forma, esta serie de estadísticas de pesca han constituido para los investigadores una gran desconocida, en especial los datos anteriores a la publicación de los primeros anuarios estadísticos. Considerando, en primer lugar, autores más o menos contemporáneos a su publicación, en Berthelot (1867) sólo se proporciona información del "Estado" de 1861. Por su parte, Ramírez y Navarrete (1905), tan sólo citan los correspondientes a los periodos siguientes: 1866-1869, 1878, 1881 (sic), 1887 (sic) y 1892, mientras que Fernández Duro (1866b), que era Secretario de la Comisión Permanente de Pesca, trata de analizar los efectos de la pesca de bous en el litoral y sólo utiliza datos de 1831 y 1861. Finalmente, Ricart (1895), menciona el de 1831, el de 1867, uno para 1887 (sic), citando a García Solá (1888), y el de 1878, comparándolos con el último, el de 1892. Es decir, parece que para los contemporáneos la serie publicada en el "Estado General de la Armada" pasó completamente desapercibida y, lógicamente, ello ha provocado también que haya pasado, en alguna medida, desapercibida para los investigadores posteriores.

En el principal trabajo publicado hasta ahora sobre fuentes estadísticas históricas del sector, el de Giraldez (1991), aunque se menciona que la serie se inicia en 1829, tan sólo hace referencia a las estadísticas que se publicaron en la "Revista de Pesca Marítima" y en los "Anuario de la Comisión Permanente de Pesca". De hecho proporciona una serie de pesca para el conjunto del Estado, iniciándola en 1883. En Burgos y Lacomba (1993) se analiza parte de la información existente en el Archivo "Álvaro de Bazán", correspondiente a la primera mitad del XIX, siendo el trabajo más significativo respecto a la localización de documentación estadística en archivos. Ocampo (2001), refiriéndose a Asturias, sólo proporciona información de algunas de las publicadas en la segunda mitad del XIX, mientras que Viruela (2000) limitándose a la actual Comunidad Valenciana tan sólo facilita información de 1831, 1861, 1883, 1889 y 1992. Por su parte, López Losa (2002), en un trabajo referido exclusivamente al País Vasco, menciona la existencia de las estadísticas de 1831, 1832, 1847, 1848, 1850, 1851, 1852, 1853, 1861 y 1883, no incorporando parte de las publicadas en el "Estado General de la Armada", ni las publicadas en "Anuario Estadístico de España" ni las dos últimas de la "Revista de Pesca Marítima"[66]. Lacomba (2006), refiriéndose a Andalucía, sólo proporciona información de los tres últimos "Estados" publicados; es decir, los correspondientes a 1883, 1889 y 1892. Y tampoco en los trabajos existentes sobre la historia de la estadística oficial española tampoco mencionan la existencia de estas estadísticas, a pesar de haber sido publicadas en los primero "Anuarios Estadísticos"[67].

---

[65] Los trabajos contemporáneos sólo mencionan los datos de los que disponemos. Tras lamentarse de la incidencia que los trastornos políticos habían tenido en la Estadística Pesquera, Salas y García Solá (1876), pág. 39, detallan el hecho de que "*en el tiempo transcurrido desde la publicación del último Anuario* - el de 1869 - *sean pocos los estados del producto de la pesca, embarcaciones y demás datos que periódicamente deben suministrarse a este centro*".

[66] Desgraciadamente López Losa (2000), pág. 249, sólo cita como fuente de sus datos del Distrito de San Sebastián, su Tesis Doctoral no publicada, por lo que es imposible dilucidar la causa de un posible error de fecha para el "Estado" correspondiente a 1848-1849, cuyo montante asigna al año "1847/48", ni el origen de un "Estado" en "1849/50", que fue publicado en la "Revista General de la Armada" y que, por tanto, permitiría disponer de una año más en la serie del "*Estado General de la Pesca*".

[67] De hecho, en ningún trabajo realizado entre la original obra de Sanz Serrano (1956) hasta la más reciente de Merediz Moreno (2004) se realiza ninguna referencia a estas estadísticas. En este trabajo, en la sección denominada "*Estadísticas agrarias y pesqueras del siglo XIX*", a pesar de asignarle esta denominación, no hace mención a ningún trabajo estadístico relativo a la pesca marítima, ni inserta en los anexos legislativos ninguna regulación sobre el particular.

La calidad de las Estadísticas de Pesca y su fiabilidad descriptiva

El "Estado General de la pesca beneficiada por los matriculados", encabezamiento más frecuente de los que aparecen en las estadísticas que estamos analizando, era elaborado a través de información trimestral remitida por los gremios de mar, como ya hemos comentado, a partir de los registros correspondientes a los arbitrios sobre capturas que éstos utilizabas para financiarse. Errores aritméticos comprobados en las primeras estadísticas y, sobre todo, los problemas derivados de una errónea agregación de las variables stock (número de embarcaciones y matriculados), además de la existencia de algunas lagunas para provincias o distritos concretos en algunos años, sobre las que hablaremos posteriormente, son las principales deficiencias detectadas, y no son atribuibles a los gremios, sino al tratamiento posterior de la información primaria[68].

Cada "Estado" distingue entre la pesca consumida en fresco y la pesca procesada. Para la primera se diferencia entre la consumida por los pescadores – por la que no tendrían que abonar arbitrios - y la consumida en la localidad de referencia. Las capturas procesadas, ya fuese en salazón o en escabeche, son las que se comercializan fuera de cada localidad del litoral, distinguiéndose entre la remitida al extranjero y la que era enviada al resto del país[69]. Por esta razón, en la mayoría de los distritos se verifica que la diferencia entre el total de la pesca capturada y la consumida en fresco coincide, lógicamente, con la suma de la pesca salada y la escabechada. Asimismo, debe coincidir también con la suma de la pesca exportada y la remitida al resto del país. No obstante existen algunos casos concretos en los que estas relaciones no se verifican. En el Departamento de Cádiz tan sólo existe algún pequeño error a lo largo de la serie, que puede ser subsanado teniendo en cuenta lo anterior, y lo mismo ocurre con el Departamento de Cartagena. No obstante, en el Departamento del El Ferrol existen bastantes deficiencias en este sentido. Unas son fáciles de subsanar, debido a que consisten, por ejemplo, a no haber incorporado en un Distrito la cantidad de pesca procesada pero si estar consignada la exportada. Otras se deben, quizás, a errores aritméticos al procesar los datos de cada población del Distrito y son de difícil solución.

Respecto a la ausencia de datos concretos, corresponden, básicamente, a la valoración monetaria de la pesca salada – muy frecuente en los distritos de Barcelona, Tarragona y Tortosa, así como en Motril, y esporádicamente en algún que otro Distrito Marítimo como Motril, Sanlúcar, Canarias, Vivero, Gijón y Santander - así como a facilitar exclusivamente información sobre pesca capturada y su valor – como ocurre en Bilbao entre 1846 y 1853 – o a una laguna absoluta, como es el caso de los distritos vascos en el "Estado" de 1845-1846.

Pero, quizás, independientemente de los errores en los que se incurría al procesar la información primaria, lo que habría que plantearse es acerca de la calidad de ésta, y de la fiabilidad de los datos facilitados por los gremios. Como hemos visto, Fernández Duro (1868) estimaba que hasta la supresión de los gremios la fiabilidad era relativamente alta, opinión

[68] El "*Estado*" de 1832, por ejemplo, estaba plagado de errores aritméticos y de agregación, hasta el punto de que mereció que por R.O. de 12 de junio de 1834 se exige a los capitanes generales de los Tercios de la Real Armada que "*dicten las prevenciones más estrechas…a fin de que cada uno de ellos remita al empezar el primer semestre en que estamos el Estado correspondiente a la suya, claro, exacto y puntual, no consintiendo por su parte la más pequeña falta a los directores de los gremios de mar y demás personas a quién toque las primeras noticias del pescado cogido, salado, escabechado, consumido en fresco, exportado al extranjero de una y otra especie, introducido en el Reino en iguales formas, fanegas de sak que se hubiesen consumido, embarcaciones y matriculados que se hayan empleado y valores de su producto*" (AMN, Mns. 447/027, pág. 117).

[69] Debe advertirse que se refiere a la pesca exportada directamente por los pescadores o comunicada a los gremios. Gran parte de las exportaciones españolas de pesca salada provienen en el XIX de pescado fresco importado desde Portugal, por lo que no deben coincidir la cifra de exportaciones dada por los "Estados" y la recogida en las estadísticas de comercio exterior.

también defendida por Ramírez y Navarrete (1905)[70]. No obstante, también en esta primera fase, mientras se mantenía la estructura gremial y la matrícula de mar, existían ciertas tasas de ocultación debidas a la desconfianza de los armadores o, incluso, a la existencia de fraudes relativos al consumo de sal que la producción podría evidenciar. A este respecto, en 1847 un funcionario de Marina comentaba las dificultades inherentes a la obtención de la información, afirmando que "no habiendo otros medios de adquirirlas fue preciso solicitarlas de los principales empresarios, únicos que llevan cuenta de sus gastos y productos, y estos, recelosos, al parecer de que la exactitud de su noticias pudiera perjudicarles atrayéndoles mayores impuestos, han disminuido siempre la cantidad de pescado cosechada y su valor, a tal punto que no vacilo en declarar que la efectiva es el doble, por lo menos"[71].

Con posterioridad, en otro contexto institucional, para paliar esta situación, con fecha 11 de julio de 1884 se establece un plan para la elaboración de una estadística general de Marina, cuyos modelos de recogida de información se publicaron en septiembre de dicho año. Poco después, a partir del primer número de la Revista de Pesca Marítima en 1885 se incorpora una sección fija denominada "Estadística, Mercado y Variedades" en cuyo preámbulo se afirma: "…la estadística no es, como vulgarmente se cree, una simple reunión de datos y número ordenados con más o menos criterio. La estadística aspira a algo más: de los grandes grupos de datos y cifras, o de los grandes número, como dicen los profesores de esta ciencia, hay que deducir principios prácticos, leyes generales, conclusiones más o menos definitivas, que puedan utilizarse convenientemente en el desarrollo, fomento y perfección de la industria a que la estadística se refiere"[72]. Estos esfuerzos son los que conducen a la publicación del "Estado General" de 1889, en el que se reconocía que era "deficiente y sin ofrecer las necesarias condiciones de exactitud" pero que, en definitiva, eran los "deducidos de los datos oficiales remitidos al Ministerio de Marina, por las autoridades de las provincias marítimas"[73]. Finalmente, en 1893 el director de la Revista de Pesca Marítima, Rafael Gutiérrez Vela fue facultado para "coleccionar los datos de pesca existentes en el Ministerio de Marina" publicando los de 1892, "únicos que hasta hoy se han podido completar"[74]. El autor mencionaba que uno de las principales defectos de las estadísticas residía en "que no se expresa por separado los resultados de la pesca de las más principales especies"[75] pero aunque reconocía que este defecto no era exclusivo de nuestro país, proponía que para mejorarlas "han de expresar en detalle, por un lado, las especies más principales objeto de la explotación, y por otro lado también los resultados que se obtienen con los procedimientos más usuales empleados para la captura"[76], estructura que fue la que se utilizó en las estadísticas que se elaboraron para 1904. De hecho, la estructura es algo diferente de la de los anteriores estados. La información se proporciona trimestralmente y se incorpora, al final, un cuadro estadístico con el número de artes de cada tipo existente en España. Por lo demás, mantiene las mismas variables que el primer "Estado".

---

[70] En Ramírez y Navararrete (1905), págs. 23-24, se afirmaba que "*hasta el año 1864 en que fueron suprimidos los Gremios de Mareantes, la conocía el Ministerio de Marina directamente con exactitud, porque como uno de los recursos que arbitraban aquellas asociaciones era el impuesto que establecían por cada arroba de pescado, los estados de producción facilitados por los gremios partían de una base conocida que presentaba bastante garantías de certeza*".
[71] AHN, Mns. 2203, fol. 81.
[72] Anexo al número de 1885 de la *Revista de Pesca Marítima*, pág. 1. En dicho número de se publicaron datos de Comercio Exterior de productos pesqueros (1877-1881) así como de la pesca de sardina (1877-1881), que habían sido publicados previamente en la última "Memoria de Industria y Legislación".
[73] Se publicaron en 1890 en la *Revista de Pesca Marítima*, págs, 167-175, 193-196 y 225-227.
[74] Gutiérrez Vela, R., (1894), pág. IV.
[75] Ídem, pág. IV.
[76] Ibídem, pág. VIII.

En definitiva, a la luz de los testimonios, la serie mantuvo ciertas dosis de coherencia y fiabilidad hasta, al menos, la extinción de los gremios y la matrícula de mar. A partir de dicho momento y hasta la consolidación de las Cofradías de Pescadores y la obligatoriedad de la subasta en primera venta en los años treinta, todas las estadísticas pesqueras adolecerían del mismo problema, dado que derivaban de declaraciones de armadores recopiladas por las Comandancias de Marina. Los investigadores actuales suelen extrapolar las críticas a la serie posterior a la desaparición de los gremios a toda la serie, lo que no parece razonable a la vista de los diferentes textos analizados[77].

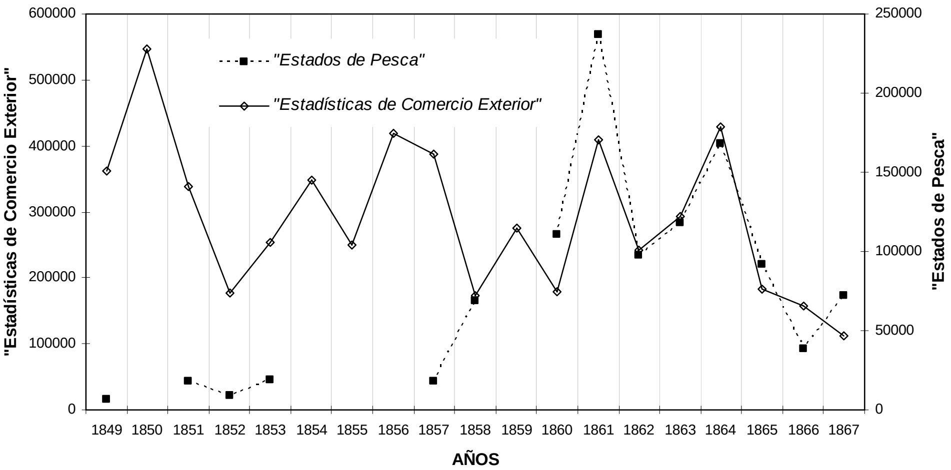


Figura 4. **Comparación de las cantidades exportadas en arrobas según los “Estados de Pesca” y la “Estadísticas de Comercio Exterior” entre 1849 y 1867**.

De cualquier forma, es posible contrastar en alguna medida la fiabilidad de los datos a través de fuentes alternativas. Por ejemplo, podemos considerar si la información recogida sobre pesca exportada resulta válida mediante su comparación con la serie de Estadísticas del Comercio Exterior[78]. En la Figura 4 hemos representado la serie de las exportaciones de los “Estados de Pesca” y las cantidades de pescado salado, ahumado o en salmuera exportado, en sus diferentes partidas, extraída de dichas estadísticas entre 1849 y 1867. Independientemente de la escala, dado que el peso en la primera de las series se refiere a peso en fresco y en la segunda al peso una vez salado, prensado y espichado[79], hay una coincidencia bastante elevada en el periodo 1858-1867, con una correlación entre ambas series de 0.88, mientras que existen fuertes discrepancias en 1849-1857. La causa de ello, que ya mencionamos anteriormente, reside en que en el primero de los periodos se producen ingentes importaciones de pesca fresca portuguesa que los fomentadores, independientemente de los gremios, salan y destinan a la exportación, beneficiándose de una política arancelaria errática que alternaba años fuertemente proteccionistas con periodos de cierto desarme, de forma que el perfil de la serie de exportaciones recoge entre 1849 y 1857 las fluctuaciones en las importaciones de

[77] Así lo manifiesta Viruela (1995), pág. 118, López Losa (2000), pág.246 o Giraldez (1991), p. 513, que se posiciona, incluso, en una postura aún más extrema, afirmando que *“no será hasta el último tercio del siglo cuando los datos comiencen a ser mínimamente coherentes*”. Diferente es la postura de Ocampo (1992), quién se postula en la dirección que estamos manteniendo.

[78] Esta serie anual se inicia en 1849 bajo el título “*Cuadro general del Comercio Exterior de España con sus posesiones ultramarinas y potencias extranjeras*” que mantiene hasta 1856, cuando toma la denominación de “*Estadística general del Comercio Exterior de España con sus posesiones de ultramar y potencias extranjeras*” con la que llega hasta 1897.

[79] Siguiendo, por ejemplo, a Herrera (1887), pág. 172, *“…para elaborar 100 kilos de sardina prensada, se necesitan 200 kilos frescas… la sardina prensada ha perdido la mitad de su peso natural y primitivo en las labores de salmuera y prensa*”.

materia prima por los fomentadores. De hecho, en dicho periodo los derechos de importación de la pesca fresca se mantienen muy elevados, situándose en unos 5 rs. vn. por arroba importada en 1854, reduciéndose posteriormente hasta situarse en un quinto de esta cifra a la vez que se endurecen las condiciones para la importación de salazones y ahumados, con la excepción del Bacalao[80].

Pero, además, esta política comercial, favorecedora de la pesca y los fomentadores nacionales, se ve acompañada por una reducción de los costes de producción a través del descenso del precio de la sal, materia prima básica para la conservación del pescado. Desde finales de la Edad Media la sal constituía un producto estancado cuyo monopolio lo detentaba la propia Corona, que aprovechaba esta situación para la obtención de fondos extraordinarios en épocas de especiales necesidades mediante la imposición de recargos sobre el precio básico, financiando así la red de caminos reales y otras obras públicas[81], el mantenimiento del Cuerpo de Milicias[82], así como los gastos ocasionados por diferentes conflictos bélicos[83], constituyendo la "renta de la sal" uno de los principales ingresos de la Corona. Desde 1782, para favorecer a los matriculados, se había establecido un "precio de gracia" a pié de fábrica de 10 rs. vn. por fanega para la sal consumida por pescadores y fomentadores a lo que se añadían los gastos de conducción desde ésta al alfolí correspondiente[84]. La situación cambia drásticamente a partir de 1823 cuando se fija un precio general de 42 rs. vn. para todos los consumidores[85]. Éste se eleva aún más en 1834[86], de forma que hasta 1854 el precio a pagar por la sal se mantiene en 52 r.s vn. por fanega consumida, con ciertas bonificaciones para los armadores, pescadores y fomentadores de la pesca, sólo para los salazones y excluyendo el salpresado, siempre y cuando su producción fuese remitida por barco a una distancia superior a las 20 leguas, lo que en realidad beneficiaba a las grandes empresas de salazón y perjudicaba los intereses de las pequeñas "chancas" y de los pescadores, cuyos volúmenes de capturas eran insuficientes para que fuese rentable su remisión a mercados alejados, mientras

---

[80] De forma simultánea a este desarme arancelario selectivo, se produce un endurecimiento de las condiciones para la importación de pescado elaborado, de forma que frente a los 0.9 rs. vn. por arroba exigidos en 1854, a partir de 1856 se imponen derechos mucho más elevados, comprendidos entre 13 y 16 rs. vn. por arroba importada, extendiéndose esta política arancelaria hasta 1869.

[81] Mediante R. D. de 10 de junio de 1761 se impone "*un sobreprecio de dos reales en fanega de sal para continuar el canal de castilla y hacer caminos rectos y sólidos en España*", manteniéndose hasta 1833.

[82] Por R. D. de 18 de noviembre de 1766 se impuso un nuevo sobreprecio de 2 rs. vn. para esta finalidad, que se elevaron a 3 rs. vn. en 1817, situación en la que se mantuvo hasta 1821.

[83] Por R. D. de 17 de noviembre de 1779 se establece un nuevo recargo de 4 rs. vn. por fanega cuya vigencia se mantuvo hasta la finalización del conflicto bélico con Gran Bretaña en 1783. Por R. D. de 17 de marzo de 1794 se impone, nuevamente, un recargo de 4 rs. vn. para financiar la Guerra de la Convención contra la República Francesa, al que se añaden otros 24 rs. vn. establecidos por R. D. de 5 de febrero de 1795, que mantuvo su vigencia hasta 1823, si bien por R. D. de 23 de enero de 1796 se redujo a 14 rs. vn.

[84] En 1782, con los recargos y sobreprecios correspondientes, el precio de la sal para los matriculados era de 19 rs. vn., mientras que el precio para el común se situaba en 30 rs. vn. Esta situación se modifica por Resolución de 23 de diciembre de 1782 mediante la que se establece un precio único para pesquerías de 10 rs. vn. por fanega a pié de fábrica, estableciendo la inaplicabilidad de los sobreprecios de milicias y caminos a los pescadores, situación inalterada hasta 1821.

[85] Durante el Trienio Liberal se "*desestanca*" la sal, se establece la libertad de comercio de la misma y por Decreto de las Cortes de 29 de junio de 1821 se establece un precio único en los alfolíes públicos de 6 rs. vn. por fanega que se elevan a 12 r.s mediante el Decreto las Cortes de 29 de junio de 1822. La restauración absolutista anula por Decreto de la Regencia de 11 de junio de 1823 todas las normas sobre el particular promulgadas por Las Cortes, y mediante R. D. de 16 de febrero de 1824 se establece el precio de 42 rs. vn. por fanega para todos los consumos.

[86] El artículo 31 del R. D. de 3 de agosto de 1834 elevó el precio de la sal a 52 rs. vn., estableciendo una bonificación del 15% para el consumo de los "fomentadores" y "pescadores" en producciones dirigidas al mercado español que se elevaba al 30% en el caso de las exportaciones. Pero ante los conflictos surgidos, la R.O. de 26/11/1835 fijó para la pesca exportada en 10 rs. vn. por fanega y 12 rs. vn. para las colonias de ultramar y la península, siempre y cuando se remitiese a más de 20 leguas del lugar de producción.

que el coste de la sal les impedía vender su producción salpresada en sus mercados tradicionales, por lo que su única alternativa era satisfacer la exigua demanda local o vender sus capturas a los grandes fomentadores, donde debían competir con las importaciones que éstos realizaban de pescado fresco portugués.

De hecho, entre 1850 y 1857 se importan anualmente unas 3.000 Tm. de pescado fresco portugués desde las aduanas onubenses y gallegas, descendiendo con posterioridad hasta situarse en una 800 Tm./año[87]. A pesar de ello, y aunque el sistema de "premios" en función de la distancia al mercado de destino se mantuvo hasta 1869, lo cierto es que la reducción del precio general de la sal en 1854 a 40 rs. vn. por fanega, cuya vigencia se extendió hasta 1868, posibilitó una cierta recuperación del sector que se detecta claramente en las estadísticas, de manera que se experimenta un sensible incremento de la producción, que pasa de las 29.100 Tm. medias anuales del periodo 1845-1853 a las 67.200 de 1854-1867[88].

Por tanto, hasta 1854 se mantuvo el precio de 52 rs. vn. por fanega reduciéndose éste en 12 rs. vn. desde dicho año y hasta 1868, con el efecto previsible sobre la producción de salazones, así como se indujo la reducción de las importaciones de salazones mediante un fuerte incremento en el arancel de los derechos aplicados sobre los mismos. Como resultado la producción del sector pesquero y la exportación de salazones por éste se duplicó en las dos décadas consideradas.

Con posterioridad, como hemos visto, los datos disponibles de los "Estados" son mucho menos frecuentes, pero permiten vislumbrar algunas características de la evolución del sector. El nivel de producción se mantendrá cercano al observado en la década de 1860, pero se comprueba un fuerte incremento de los trabajadores del sector, que crecen desde un promedio anual de 37.000 matriculados entre 1864-1867 a los 66.000 pescadores que recogen los datos de 1878 y 1883. Éste incremento se debe, fundamentalmente, a la desaparición de la "matrícula de mar" y subsiguiente liberalización del sector y, sobre todo, a que ambas magnitudes no resultan comparables, dado que durante todo el siglo XIX se mantienen vigentes las observaciones que realizamos acerca del empleo pesquero en las "inspecciones de matrículas"; es decir, la presencia de "terrestres" en labores pesqueras siempre y cuando su función no implicase realizar su trabajo embarcados[89].

Pero en la época que nos ocupa - sin una causa aparente de escasez de matriculados por el excesivo reclutamiento de la Real Armada como ocurrió a finales del XVIII y principios del XIX – se van produciendo continuas aprobaciones de licencias y permisos en diferentes

---

[87] Lo injusto de la medida era evidenciado en diferentes memorias, como por ejemplo en una datada en 1839 en la que se afirmaba que "*…no habiendo las leguas marcadas por la Ley para el abono de los cuarenta reales por fanega consumida en la salazón, la han de pagar aquellos empresarios a cincuenta y dos reales, y en este caso ¿qué ventajas van a conseguir aquellos fomentadores de pescas? La obtendrán los pueblos a los que quepa la suerte de hallarse fuera del radio de las veinte leguas, los que podrán en el mismo mercado, darla a precios más cómodos, y lucrar; y se perderán los que estando dentro de dicho radio, hayan de seguir el precio de la plaza"* [Miravent, (1850), págs. 79-80].

[88] El R.D. de 1 de abril de 1854, en su artículo 1º, fijó el precio de la sal en 40 rs. vn. por fanega, manteniendo las bonificaciones establecidas anteriormente para pescadores y fomentadores. El R. D. de 4 de marzo de 1869 fijó un precio único para la industria y los fomentadores de 10 rs. vn. por quintal. Finalmente, por Decreto de Cortes Constituyentes de 16 de junio de 1869 se establece el definitivo desestanco de la sal y la completa liberalización de su producción y comercio.

[89] En 1847, en las respuestas a un interrogatorio sobre la pesca, el Comandante de Marina de la Provincia de Huelva declaraba que en ésta se ejercitaban "*…de setecientos cincuenta a ochocientos matriculados y seiscientos terrestres*" AMN, Mns. 2203, fol. 96. La primera cifra coincide con la de los "Estados", por lo que éstos, hasta la desaparición de la Matrícula de Mar, están subvalorados.

puntos del litoral para que los “terrestres” puedan incorporarse a la actividad pesquera[90], llegándose hasta el punto de permitir la incorporación de matriculados portugueses en las actividades pesqueras y marítimas en Huelva[91] o a la de autorizar el ejercicio de la pesca, de forma general, a todos los matriculados inhábiles[92].

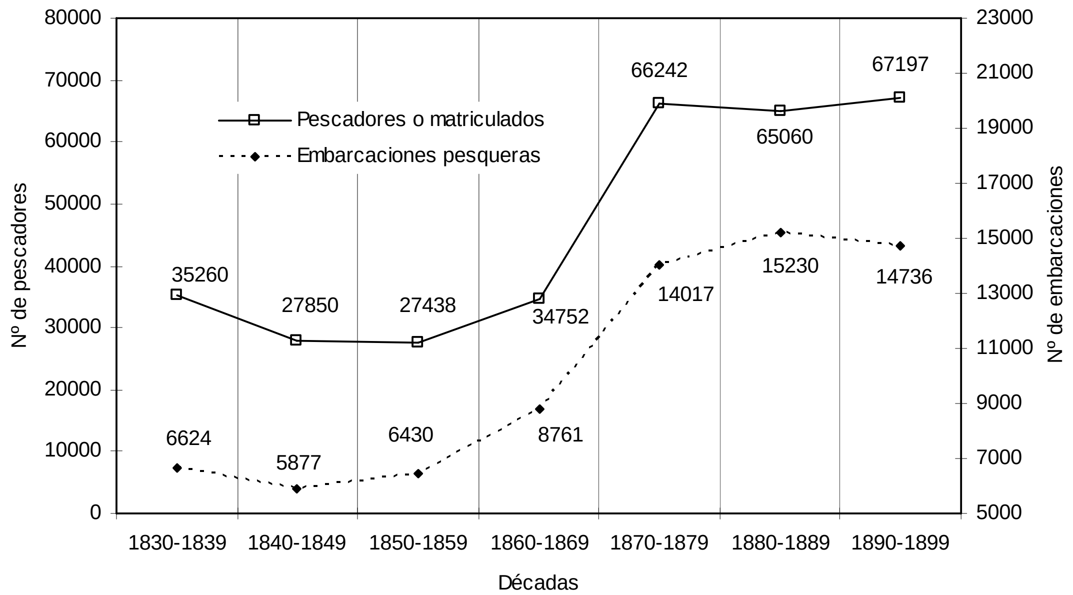


Figura 5. **Evolución decenal del número medio de embarcaciones y pescadores para el conjunto nacional según los “Estados” del siglo XIX**

Todo ello resulta una prueba más que evidente de la expansión que estaba experimentando la actividad pesquera y coincide con lo que se deduce de los “Estados de Pesca”. En efecto, debe entenderse que hasta la definitiva desaparición de la Matrícula de Mar la cifra de empleo de los “Estados” estaría subvalorada, dado que excluye al gran número de terrestres dedicados, legal o ilegalmente a esta actividad.

Con posterioridad a la desaparición de esta institución se produce entre 1867 y 1878 un fuerte incremento del empleo recogido en las estadísticas, que pasa de 37.537 matriculados a 66.242 pescadores – un 76.4% - mientras que el número de embarcaciones experimenta un incremento sensiblemente inferior, pasando de 10.216 embarcaciones en 1867 a 14.017 barcos en 1878; es decir, un 37% de incremento.

Es lógico pensar, por tanto, que el cambio institucional experimentado es la causa principal del fuerte incremento del empleo, haciendo surgir de las penumbras estadísticas al contingente de los terrestres que se empleaban en la pesca, que pueden ser estimados en algo

[90] La primera norma al respecto surge en los meses previos a la batalla de Trafalgar, describiéndose en la Real Cédula de 31 de marzo de 1805 la situación claramente: “*considerando que con motivo de la presente guerra tendrán que salir de los puertos todos los matriculados útiles, y que quedarán por consiguiente sin ejercicio los mencionados barcos y aparejos, los pueblos sin pescados, las familias de la gente de mar sin arbitrios para subsistir*”. Pero en ausencia de grandes reclutamientos, se establecían en la época que ahora nos ocupa nuevas licencias, mediante la R.O. de 1 de diciembre de 1845, derogada parcialmente por la R.O. de 15 de abril de 1852, o se autoriza nuevamente en El Ferrol por R.O. de 30 de mayo de 1860, en la costa de Huelva por R. O. de 10 de octubre de 1861, o en Santander por Real Orden de 27 de febrero de 1867.

[91] Por R. O. de 22 de noviembre de 1861 actualizada mediante la R.O. de 11 de noviembre de 1864.

[92] Temporalmente por R.O. 15 de octubre de 1861 y definitivamente en la R.O. de 22 de agosto de 1863.

más de 11.000 trabajadores en todo el litoral, obtenidos al aplicar a la flota de 1867 de cada distrito el número medio de tripulantes por embarcación existentes en 1878.

Por tanto, y en definitiva, los "Estados" ponen de manifiesto la expansión de la actividad pesquera resultante de la liberalización del sector. La desaparición de los gremios en 1863, el desestanco de la sal en 1869 y la supresión de la matrícula de mar en 1874, provocan la afluencia de capitales al sector y su progresiva modernización. A pesar de las deficiencias, de las lagunas y de sus limitaciones metodológicas, los "Estados" constituyen una fuente fundamental para reconocer la evolución del sector pesquero español a lo largo del siglo XIX. En los anexos a este trabajo hemos incorporado los datos correspondientes a pesca capturada, valor, embarcaciones y empleo en la pesca según los "Estados", tanto las series originales como las revisadas en las que se han completado lagunas –asignándole a la provincia correspondiente los datos del año más próximo temporalmente – y se han corregido algunos errores de agregación, dado que como se ha mencionado, a través de estados semestrales, trimestrales o mensuales se obtenía el total anual y, lógicamente, mientras que el total de las magnitudes flujo se obtenía por agregación, para las variables stock – número de embarcaciones y número de tripulantes – debía obtenerse una media o la moda de los diferentes estados estacionales, pero era frecuente que figurasen la suma de los estados mensuales , trimestrales o semestrales.

Otras estadísticas pesqueras del siglo XIX

Los "Estados" no fueron las únicas estadísticas pesqueras de carácter periódico confeccionadas a lo largo del siglo XIX, sino que coexistieron con otras investigaciones concretas que trataban de aportar información sobre pesquerías específicas tales como las almadrabas, la pesca del coral o las actividades extractivas desarrolladas en las aguas del Mar Menor.

Las almadrabas, grandes artes de red dirigidas a la captura de atún rojo y otras especies afines[93], habían sido explotadas en régimen de monopolio por concesión Real desde la Edad Media por la Casa Ducal de Medina Sidonia en la costa andaluza, mientras que en el Tercio de Levante unas eran explotadas por los gremios, otras por casas señoriales y otras por empresarios que habían obtenido su concesión[94]. Curiosamente, la primera recopilación de datos estadísticos de pesca, con un horizonte temporal que abarcaba más de dos siglos, se realiza como resultado de una Memoria que Martín Sarmiento elabora para el Duque de Medina Sidonia en 1757. Explotando los legajos del archivo de la Casa Ducal, Sarmiento sistematiza las capturas anuales en número de atunes desde 1525 a 1756 de las siete almadrabas que entre dichas fechas fueron caladas en uno u otro periodo y cuyo perfil, por constituir las más antiguas estadísticas pesqueras existentes, reproducimos en la Figura 5[95].

[93] Existían varios tipos de almadrabas, siendo los más relevantes la de Tiro o Monteleva, arte playero de tiro que no era más que una jábega de dimensiones extraordinarias o las de Buche, precedente directo de las actuales, que consistía en un circo de redes que se mantenía calado durante toda la temporada. Una discusión sobre sus tipos problemática puede verse en García del Hoyo, J.J., (2002), pág. 32-38.

[94] En Sañez Reguart, A., (1791), Tomo I, págs. 6-70, se mencionan unas 12-13 almadrabas en Levante, es decir, desde Almería a la frontera francesa y otras 4-5 caladas habitualmente y pertenecientes a la Casa Ducal mencionadas en la costa comprendida entre Tarifa y el Guadiana. Lógicamente, dependiendo de diferentes factores se calaban más o menos almadrabas, pero hasta su liberalización parece que la localización de las mismas era fija.

[95] La memoria de Fray Martín Sarmiento, titulada "*De los atunes y sus transmigraciones*" se encuentran en un manuscrito depositado en el Museo Provincial de Pontevedra. Se han realizado dos transcripciones completas del

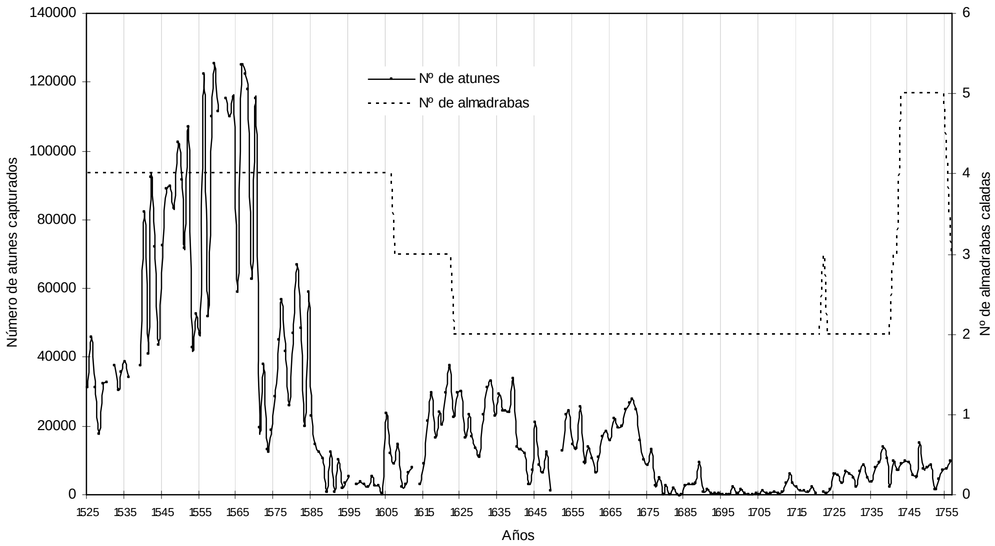


Figura 6. **Evolución de las capturas de atunes y del número de Almadrabas caladas por la Casa de Medina Sidonia entre 1525 y 1756 según Martín de Sarmiento**.

La Ordenanza de Matrículas de 1802 ya había planteado la posibilidad de eliminar los privilegios señoriales sobre estos artes, pero son las Cortes de Cádiz las primeras en mencionar expresamente esta eventualidad[96], provocando la aparición de las primeras almadrabas caladas por los gremios, de forma que tras la Guerra de la Independencia y la restauración de Fernando VII, el Ministro Vázquez Figueroa promueve su extinción: "…queden abolidos para siembre todos los privilegios exclusivos relativos a la pesca concedidos a particulares o corporaciones de cualquier clase que sean, y declaro la facultad de emplearse en ella a todos mis vasallos, con la condición de alistarse a las matrículas de mar"[97]. Es entonces cuando se produce la multiplicación de los pesqueros, especialmente en el Golfo de Cádiz y el Estrecho de Gibraltar, donde frente a las tres almadrabas caladas por el Marqués de Villafranca en 1804 se calan 11 almadrabas en 1817[98]. Pero la actividad entre dicho año y 1828 se había reducido drásticamente a causa de las limitaciones impuestas a los "buches" y, sobre todo, por la persistencia de años de pésimas capturas, de forma que entre 1823 y 1827 sólo fue calada la almadraba de Zahara. En 1828, se publican los reglamentos definitivos para la explotación de las almadrabas, por R.O. de 22 de agosto para las del Tercio de Levante y por R. O. de 24 de septiembre para las del Tercio de Poniente, incluyéndose en

manuscrito, la preparada por F. López Capont (1997), en edición facsímil, y la de J.L. Pensado (1992), aunque esta última no incluye los datos estadísticos.

[96] Según el Decreto de Cortes de 6 de agosto de 1811 "*aboliendo los privilegios exclusivos, privativos y prohibitivos que tengan el mismo origen de señorío, como son los de caza, pesca, etc*".

[97] R.D. de 20 de febrero de 1817.

[98] Por Oficio de 22 de mayo de 1817 fue nombrado el Capitán de Navío José Agustín de Lobatón para realizar un análisis de la situación de las almadrabas y proponer un Reglamento provisional para sus uso, que remitido por éste el 17 de julio de 1817 (AMN, Mns. 443, págs. 72-75), fue aprobado por R.O. de 2 de marzo de 1818, donde se atribuía a los gremios de matriculados la titularidad de las almadrabas. Es, precisamente, en estos informes en los que se contienen los primeros estados de pesca de las almadrabas de poniente, con mención a los empleos generados y el valor de las rentas obtenidas, y corresponden a 1817. Posteriormente, en un informe fechado en 1 de agosto de 1818 sobre los resultados de una pésima temporada, se incorpora el "estado" correspondiente a dicho año en el que se documenta, por primera vez, el número de atunes capturados, que ascendió a tan sólo 1667 atunes en las seis almadrabas que operaron (AMN, Mns. 443. 205-206).

ambas que “las pesquerías de almadrabas quedan consignadas en propiedad a los gremios”.[99] Los gremios, que estaban obligados a disponer de todo el material para su calamento, podrían explotar directamente la almadraba o bien, mediante subasta, conceder su explotación por un año a particulares cuando el gremio aportase los materiales o por dos años si éstos fuesen aportados por el adjudicatario. En caso de no haberse adjudicado por subasta, el gremio estaba obligado a calar el arte. Esta posibilidad planteó no pocas corruptelas, siendo frecuente que los gremios pactasen con el posible adjudicatario el que la subasta quedase sin remate, dado que entonces sólo tenían que abonar a la Marina “un real de vellón por cada arroba de cincuenta libras de pescado que coja”. Se reabre entonces, además, el conflicto entre las tradicionales almadrabas de tiro caladas por los gremios - intensivas en mano de obra - y las más sofisticadas almadrabas de buche – intensivas en capital - logrando finalmente los primeros la prohibición de éstas entre Cádiz y Tarifa a partir de 1837, que se extendió hasta 1844, se renovó en 1847 y se derogó, definitivamente, en 1888[100].

Es en este contexto cuando se hace evidente la necesidad de que se disponga de información suficiente acerca de las capturas y medios de las almadrabas[101]. Por ello mediante R.O. de 16 de mayo de 1843 se establece que “los Comandantes Generales de los Departamentos den anualmente, el 15 de junio, un estado general de las almadrabas existentes en la comprensión de los suyos respectivos”. En el artículo 7º de dicha norma se establecía la estructura del “Estado particular de Pesca” en el que se debía detallar para cada almadraba la cantidad del arriendo, la cantidad percibida por la Marina, el número de atunes y bonitos capturados, el número de otros peces, el valor del pescado en primera venta, la cantidad de pescado destinada a salazón, la sal consumida, las cantidades exportadas, así como el detalle de los gastos realizados por la almadraba, las ganancias o perdidas y la gente de mar empleada. Igualmente, en virtud del artículo 6º de dicha disposición, debería facilitarse la fecha de creación de la almadraba, la fecha del último remate, la cantidad establecida para el arriendo, el nombre del arrendatario y el producto quinquenal de la almadraba. Con esta información se trataba de evitar el frecuente fraude en la subasta y ajustar el precio del arriendo a la rentabilidad potencial de la almadraba.

Los “Estados” se publicaron en diferentes ocasiones, tanto en la Gaceta de Madrid, en los dos Anuarios de la Comisión Permanente de Pesca, en las tres Memorias sobre la Industria y Legislación de Pesca y en la Revista de Pesca Marítima[102]. Una cuestión relevante sobre las estadísticas de almadrabas reside en el hecho de que el producto de éstas no está incluido en la serie de los “Estados de Pesca” que han sido analizados anteriormente, por lo que esta

[99] No obstante este mismo precepto había sido ya promulgado por R.O. de 22 de febrero de 1828 como resultado de las propuestas realizadas en la revista de inspección de matrículas anterior.

[100] El Duque de Medina Sidonia calaba un buche en Huelva desde 1740. En la última década del XVIII intentó extender esta modalidad a Conil y Zahara encontrando una fuerte oposición por parte de los matriculados. Una vez eliminados los privilegios, la presión gremial consigue en repetidas ocasiones la prohibición de los buches; R.O. de15 de mayo de 1817, hasta 1818, R.O. de 25 de octubre de 1819, hasta 1828 con la excepción concretada para 1825 por R.O. de 4 de diciembre. La prohibición más duradera fue la de 1837, aunque en realidad se inició su tramitación en 1836, que se extendió hasta 1844, cuando por R.D. de 14 de febrero se reestablecen las almadrabas de buche de Zahara, Conil y Punta de la Isla, volviendo a prohibirse por R.D. de 16 de junio de 1847, extendiéndose hasta su definitiva autorización por Ley de 11 de mayo de 1888.

[101] De hecho, resulta sintomático que el “*Dictamen de la Comisión de Marina sobre el uso del arte de Almadraba de Buche en la punta de la Isla de San Fernando*” (G.M. de mayo de 1837) se limite a reproducir un párrafo en el que Sañez Reguart (1791), pág. 50, en el que éste realizaba un análisis de la serie de capturas recopilada por Martín Sarmiento para el periodo 1525-1570.

[102] También se reproducen alguno de estos “*Estados*” parcialmente, los comprendidos entre 1850 y 1860, en Fernández Duro, C., (1866a). La mayor parte de estos “*Estados*” se encuentran depositados en los legajos 2134, 2142 y 2150 del AHMAB y en los expedientes de cada almadraba depositados en el Archivo Central de la Administración en Alcalá de Henares.

producción debe agregarse a la de los anteriores para poder disponer de la producción completa del sector. Por ejemplo, y en lo que respecta a los distritos marítimos del Golfo de Cádiz, la producción de las almadrabas suponía un promedio del 8-10% de las capturas totales entre 1829 y 1850.

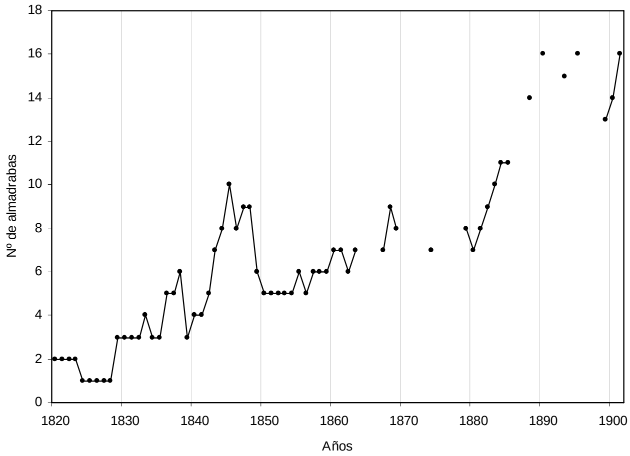


Figura 7. **Evolución del número de almadrabas caladas en el Golfo de Cádiz desde la supresión de los privilegios señoriales.**

En las dos décadas siguientes se mantuvo en torno al 20% de la producción pesquera, creciendo en las dos décadas finales del siglo XIX hasta situarse en torno al 40% de ésta. La obligatoriedad de estas declaraciones se mantuvo a lo largo de los diferentes reglamentos que se aprobaron tanto en el siglo XIX como en el XX, de forma que es posible reconstruir una serie más o menos homogénea de la estructura del sector almadrabero hasta la actualidad. De hecho, el artículo 39 del “Reglamento para Gobierno y Disfrute de las Almadrabas” de 1866, aprobado por R. O. de 2 de junio de 1866, establecía que “los Capitanes Generales remitirán al Ministerio de Marina anualmente, el 1º de diciembre, un estado que comprenda todas las almadrabas del Departamento y costa de África, expresando el año de la fundación, la cantidad en que hayan sido rematadas y el número de gente de mar empleada, y si los contratistas o concesionarios facilitasen voluntariamente datos sobre el pescado cogido y beneficiado en fresco y salado, se añadirán las casillas correspondientes”. Esta última matización, la voluntariedad de aportar la información sobre capturas provocó un progresivo deterioro de la validez de los “estados”, de forma, que es extraño encontrar, en los publicados correspondientes a las dos últimas décadas del siglo XIX, mención alguna a las cantidades capturadas[103].

Desgraciadamente, este mismo precepto fue incorporado al Reglamento de 1899[104], que derogó al anterior, cuyo artículo 40 prácticamente reprodujo el que hasta entonces estaba vigente, manteniendo la voluntariedad de las declaraciones de capturas y de la pesca destinada a la transformación. Las deficiencias del anterior Reglamento obligaron bien pronto a su sustitución, de forma que por R.D. de 9 de julio de 1908 se aprueba el nuevo Reglamento (G.M. nº 203, de 21 de julio, págs. 314-317) en el que se incorpora la obligatoriedad de la declaración de capturas en su artículo 32: “El arrendatario entregará al Ayudante de Marina,

[103] Esto ocurre con todos los publicados en la *Revista de Pesca Marítima*, correspondientes a 1889, 1892 y 1898, en los que tan sólo se incorporan los datos de obligatorio cumplimiento. No obstante, en los legajos donde se contienen los expedientes de las sucesivas subastas depositados en el Archivo General de la Administración de Alcalá de Henares hemos comprobado que se incluyen gran parte de las declaraciones de capturas, por lo que una investigación a realizar es la reconstrucción de la serie.
[104] Publicado por R. D. de 5 de abril de 1899 (G. M. nº 97, de 7 de abril de 1899, págs. 59-61.

dentro del mes siguiente al término oficial de cada temporada de pesca, un estado según el adjunto modelo. Siendo el objeto de estos estados el formar la estadística se reputarán como documentos oficiales y, por tanto, cualquier alteración de la verdad que pase de los límites racionales de equivocación será denunciada en los tribunales”[105]. No será, por tanto, hasta bien entrado el siglo XX, cuando se obligue de manera definitiva a los arrendatarios de las almadrabas a comunicar toda la información, consolidando el sistema de recogida de la información[106].

Otras estadísticas de menor entidad, pero también con cierto carácter de permanencia en el tiempo, fueron confeccionadas en la segunda mitad del siglo XIX. Por ejemplo, a partir de la promulgación del Reglamento para la pesca en la albufera del mar menor, por R.D. de 6 de mayo de 1879, comienzan a realizarse unos estados anuales sobre esta pesquería, que son publicados en los medios mencionados anteriormente. Su estructura, definida en el artículo 53 del reglamento, detallaba el producto de la pesca para cada encañizada y, lo que suponía una novedad en las estadísticas pesqueras de la época, distinguiendo cada una de las especies capturadas (anguilas, boquerón, dorada, lisas, sardina, y diferentes tipos de mugílidos) y de las huevas extraídas. Su elaboración se extendió hasta, al menos, finales de la década de 1890, cuando se publican los últimos estados en la Revista de Pesca Marítima.

Asimismo se realizaron resúmenes estadísticos sobre pesquerías concretas respondiendo a requerimiento de necesidades la pesca del coral y esponjas, sobre la pesca del bou[107] o sobre la pesca de la sardina. Esta última fue la de mayor calado de las de este grupo, dado que respondía a un cuestionario remitido a todos los puertos del litoral y permitió realizar una serie para el periodo 1877-1881, tanto de las capturas como de su procesado en salazones y conservas así como del empleo generado[108].

Conclusiones

Cuando en 1905 se celebra el Congreso Internacional de Pesca y Piscicultura en Viena, los delegados españoles acuden desesperanzados. La organización del Congreso había remitido un extenso cuestionario integrado por casi cien preguntas diferentes a las que los países asistentes debían tratar de responder en sus informes. Casi dos tercios de éstas eran relativas a la forma de confeccionar las estadísticas pesqueras, la estructura organizativa del sistema de recogida, periodicidad, fiabilidad, difusión, y un largo etcétera al que nuestros delegados muy difícilmente podían contestar. En su exposición preliminar al Ministro de Marina sobre el informe que iban a presentar declaran que “han tropezado con escollos difíciles de vencer por la falta de noticias oficiales completas y recientes para satisfacer, pregunta por pregunta, todas y cada una de las que contienen los referidos cuestionarios, sin hacer resaltar dicha falta un aducir excusas para omisiones que pudieran hacer suponer deficiencias en la organización o

[105] Estos modelos se mantendrán prácticamente sin alteración hasta finales de la década de 1960 y sus resúmenes son los que se publicaron en los anuarios “*Estadísticas de Pesca*” entre 1940 y 1970.

[106] Dado que las capturas de las empresas almadraberas o bien se han dirigido tradicionalmente a las fábricas de salazón y conservas o, actualmente, a la exportación, no han sido subastadas en lonja, por lo que, teniendo en cuenta que las subastas en lonja han constituido y constituyen aún la base para la confección de las estadísticas pesqueras, la producción de esta actividad siempre ha tenido una consideración diferente en el sistema estadístico pesquero español y, en los últimos años, en el andaluz.

[107] Se publicó un resumen por distrito marítimo con detalle de la estructura del sector, de la producción y de la rentabilidad de los artes en Fernández Duro, C., (1868), págs.73-75, y respondía a una investigación concreta, planteada en Fernández Duro, C., (1866b).

[108] Se incorporó como anexo en Gutiérrez Vela (1886) y también en el número de 1885 de la *Revista de Pesca Marítima*.

en los servicios del Ministerio de Marina"[109], dado que "desde la publicación del último Anuario por la Comisión Central de Pesca en 1892 (sic), no se encuentra impresa ninguna recopilación o estudio oficial de conjunto que permita reunir los datos necesarios para dar idea exacta del estado actual de las industrias de la pesca en España, y aún cuando esa Dirección General nos ha facilitado los trabajos realizados por el Auxiliar del Negociado de Pesca, D. Ángel Pardo, recopilando los datos de las informaciones estadísticas mandadas verificar por Real Orden de 25 de junio de 1904"[110]. Difícil misión la suya con los medios disponibles. En realidad confeccionaron, con prisas y con múltiples errores, un escueto informe en el que compararon los datos de los "Estados" de 1892 y 1878 con los que el Ministerio disponía relativos a 1904, dado que así contestaban cuestionario "sin gran riesgo de alterar la verdad". En el informe se confunden fechas de publicación de las estadísticas, se alteran datos y los que se presentaron resultaron de una gran pobreza. No obstante los delegados estaban satisfechos de su trabajo hasta el punto de informar al Ministro de que "presentamos nuestra Monografía de la pesca marítima en España, aprobada por V.E., que fue leída y ampliada verbalmente en sesión pública por el señor Navarrete, recibida con aplauso por los congresistas y agradecida, muy especialmente, por el Presidente"[111]. Tan magna proeza, salir supuestamente airosos de tan problemático trance, mereció recompensa en la forma de sendas cruces de 2ª clase al mérito Naval "con distintivo blanco sin pensión" a ambos delegados.

La "Monografía" se publicó, pero el gobierno comenzó a darse cuenta del abandono en el que se mantenía al sector pesquero, y de que había quedado en evidencia frente a la mayoría de los países europeos. Un cuarto de siglo después, con motivo de la publicación de una extensa investigación estadística, Odón del Buen, Director y fundador del Instituto Español de Oceanografía, que había sido Catedrático en las universidades de La Sorbona y de Barcelona, comentaba el incidente del Congreso, con cierta vergüenza: "No existía en nuestro país estadística de pesca; la conciencia general, fuera de España, condenaba esa desidia inconcebible; se llegó a decir en una obra clásica acerca de los clupéidos editada por el Consejo Internacional para la Explotación del Mar, refiriéndose a la estadística presentada por los Delegados españoles al Congreso Internacional de Pesca de Viena (1905), y al Anuario de Pesca, que publicaba el Ministerio de Marina, que no contenían los únicos datos que tenían utilidad. Puedo, además, afirmar con conocimiento de causa que los datos de aquellas estadísticas eran falsos"[112].

Como hemos visto, debido en gran parte a los cambios institucionales producidos, el sistema estadístico pesquero, que se había desarrollado entre 1830 y 1845 se encontraba al finalizar el siglo muy mermado; no se realizaban los trabajos con la periodicidad exigida, no se remitía la información desde las Ayudantías de Marina, y lo que era aún peor, no existía una estructura administrativa dedicada a su elaboración. No es de extrañar que éste no resista la comparación con los de los principales países de nuestro entorno, en los que las estadísticas pesqueras, en las últimas décadas del siglo XIX, se habían convertido en un firme apoyo para la investigación biológica aplicada.

En Francia, el nacimiento de las primeras estadísticas de pesca es bastante más tardío. De hecho, cuando Moreau de Jonnes publica en 1856 su "Statistique de l'Industrie de la France" apenas dispone de alguna burda estimación de la producción global y de la flota pesquera existente, teniendo que conformarse prácticamente con algunos datos de comercio exterior y

109 Véase Ramírez F. y A. Navarrete, (1905), pág. 12.
110 Ídem, pág. 13.
111 Ibidem, pág. 39.
112 Véase Del Buen, O., (1929), pág. 2.

datos de consumo de pescado en la ciudad de Paris[113]. Las primeras estadísticas pesqueras de carácter general no se publican hasta 1868 y su contenido, aunque no difería demasiado de las realizadas en España, superaba a éstas a causa de para cada distrito – quatier – de la costa detallaba cada magnitud – incluso el número de barcos y de tripulantes - en relación a cada una de las especies capturadas, aunque en estas primeras estadísticas sean frecuentes las lagunas y de que no en todos los distritos se facilitase dicho nivel de desagregación[114]. Por Circular de 9 de febrero de 1886 se define una nueva estructura a estas estadísticas, presentándose resultados generales y resúmenes por especies, estructura de la flota, establecimientos de pesca e, incluso, pescadores extranjeros que operaban en aguas francesas, acompañado todo ello por una memoria descriptiva que sustituye ventajosamente las anotaciones marginales de la anterior etapa. Las Statistiques des pêches maritimes se mantendrán de forma ininterrumpida hasta la actualidad.

En Inglaterra y Gales se crea en 1885 el Departamento de Pesquerías del Board of Trade que comienza a recopilar y difundir información sistemática sobre pesquerías desde 1886, excluyendo la pesca fluvial y el salmón, incorporando datos de pesca capturada y valor de la misma para diferentes puertos, sin incorporar información sobre flota pesquera y empleo en el sector[115]. Desde 1905 estas estadísticas comienzan a elaborar mensualmente y se amplia la desagregación espacial y de las especies capturadas[116].

En Escocia se crea en 1882 el Fishery Board que pronto comienza a confeccionar estadísticas bastante completas para la época, que superaban las elaboradas para el resto del Reino Unido y permitieron la realización de incipientes trabajos científicos sobre los recursos del Mar del Norte[117]. Se realizaban con periodicidad mensual para cada distrito y facilitaban el peso y valor de los desembarcos para cada especie, incorporando, además, información sobre costes, pero no detallaban estadísticas de flota ni del nivel de empleo. De hecho, en 1883 se había constituido una Comisión Real para el análisis de las pesquería de arrastre en dicho caladero y entre sus recomendaciones destaca la de recopilar estadísticas de captura y realizar un experimento sobre los efectos de los barcos de arrastre[118].

Como vemos, las estadísticas pesqueras de los países de nuestro entorno inmediato son bastante posteriores a las primeras estadísticas pesqueras españolas, pero una vez que comienzan a desarrollarse, son de mayor calidad y con niveles de desagregación más altos que las que en la misma época se estaban desarrollando en España. De hecho, mientras se desarrolla la estadística pesquera en Europa, en España se produce el deterioro de la misma. Pero, además, las publicadas desde 1880 en nuestro país, siguen la misma estructura de las que se habían desarrollado cincuenta años antes, totalmente obsoletas e inutilizables desde un punto de vista científico.

---

[113] Véase Moreau de Jonnes, M.A., (1856), págs. 296-300.

[114] Estas primeras estadísticas pesqueras francesas, correspondientes a los años 1865 y 1866, se publicaron en la *Revue Marítime et Colonial* de febrero de 1868. En su presentación se declaraba que en otros países como Inglaterra, Escocia o Noruega se publicaba cada año una estadística de la pesca marítima comprendiendo "*toutes les informations intéressant cette industrie*". Y añadía que "*aucune publication du même genre n'a, jusqu'à présent, été faite en France*" (pág. 389).

[115] No obstante estas primeras estadísticas eran bastante deficientes en cuanto a la agregación de especies. Como afirmaba Cunningham (1895), pág. 56, que eran "s*o irregular that it is quite imposible to understand them because of changes in the method of classifying and estimating the fish*".

[116] Un análisis de estas primeras estadísticas inglesas, especialmente de las relativas a la flota de arrastre, se encuentra en Garstang, W., (1900), incluyendo algunas de las series disponibles desde 1886.

[117] Una excelente descripción de estos primeros trabajos estadísticos y de sus aplicaciones científicas se encuentra en Fulton, T.W., (1889).

[118] Véase, por ejemplo, Smith, T.D., (1994), pág. 87.

Lo más triste de la situación es que desde hacía bastante tiempo se conocían perfectamente cuáles eran las exigencias mínimas que debían imponerse a las estadísticas pesqueras. Por ejemplo, en el primer manual de Estadística publicado en nuestro país, el del portugués Sampaio (1841), se detallaba claramente cuál debía de ser la estructura de unas buenas estadísticas pesqueras, que debían incluir “las especies, cantidad y precio del producto anual líquido y neto; el número de personas que en ellas se emplean exclusiva o simultáneamente con otras profesiones; sus salarios; modo de hacer la vida, carácter y civilización; facilidad o embarazos que se experimentan en esta industria, número de embarcaciones mayores y menores que demandan, consumo de los productos, ….”[119]. Poco después, José María Ibáñez, Catedrático de Estadística de la Matritense, era mucho más preciso en sus recomendaciones, detallando el contenido de lo que deberían de ser unas estadísticas de la pesca que para él debería detallar para cada especie un conjunto de variables análogo al que los “Estados” recogían así como facilitar información adicional sobre costes, importaciones, etc[120]. De haberse seguido sus indicaciones nuestras estadísticas pesqueras habrían estado al nivel de las de otros países cuarenta años antes.

Pero no sólo existían tratados teóricos al respecto, sino que en la Revista de Pesca Marítima se reproducían habitualmente tanto las estadísticas francesas como las escocesas o noruegas, por lo que la comparación era obvia.

En definitiva, los “Estados de Pesca” realizados entre 1829 y 1892, fruto de la Marina ilustrada, complementados por otras estadísticas como la de almadrabas, constituyeron una innovación satisfactoria en el contexto de los países de nuestro entorno. La dejadez y el atraso que nuestro país arrastró durante el siglo XIX no sólo permitieron su transformación en unas verdaderas estadísticas pesqueras, sino que incluso perdió en gran medida su fiabilidad. Como en otros aspectos, la estadística pesquera tardaría veinte o treinta años, hasta 1914, en facilitar información comparable con la que se desarrollaba en otros países no sería capaz de situarse al nivel de otros países.

---

[119] Véase Sampaio, A.P.F. de, (1841), pág. 46.

[120] De hecho Ibáñez, J.M. (1844), pág. 225 y 226, en el Tomo I de su “*Tratado*”, establece el contenido que, a su entender, debería tener toda investigación estadística sobre la pesca marítima, mientras que en el Tomo II, publicado al año siguiente, aporta, en los modelos 18 y 19, págs. 316-319, un formato concreto de los cuadros que deberían utilizarse en estas investigaciones.

Bibliografía

AMORÓS, N., (1925), Obras científicas, profesionales y literarias, Madrid: Editorial V. Campo- Radio, 412 págs.

BERTHELOT, M.S., (1867), Exploración de la Costa Meridional de España. Cádiz: Imprenta y Litografía de la Revista Médica, 82 págs. Se trata, en realidad, de la traducción del capítulo dedicado a España de su libro, Études sur les pêches maritimes dans la Méditerranée et l'Océan, publicado en París en 1868.

BUEN, O. del, (1929), "Algunas consideraciones previas a la Estadística de la Pesca Marítima en España (Año 1928)", Notas y resúmenes del Instituto Español de Oceanografía, Serie II, nº 34, págs. 1-4.

BURGOS MADROÑERO, M., (1992), "La matrícula de mar y la pesca en las provincias de Marina de Sevilla y Ayamonte en los siglos XVIII y XIX", en XI Congreso de Profesores Investigadores, Palos de la Frontera 1992. Córdoba: Asociación de Profesores de Geografía e Historia de Bachillerato de Andalucía, págs. 297-312.

BURGOS MADROÑERO, M., (2003), Hombres de mar, pesca y embarcaciones en Andalucía. La Matrícula de Mar en los siglos XVIII y XIX (1700-1850), Sevilla: Servicio de Publicaciones de la Consejería de Agricultura y Pesca, 291 págs.

BURGOS MADROÑERO, M. Y J.A. LACOMBA, (1993), "El sector pesquero en Andalucía en la primera mitad del siglo XIX. Una Aproximación", Revista de Estudios Regionales. 35, págs. 15-50.

CARRERAS, M. Y J.M. PIERNAS, (1873), Tratado Elemental de Estadística, Madrid: Imprenta y Librería de Miguel Guijarro, 330 págs.

COLL, S. Y J.L. FORTEA, (1995), Guía de Fuentes Cuantitativas para la Historia Económica de España, Vol. I. Recursos y Sectores Productivos, Madrid: Banco de España, 166 págs.

CORROZA, C., (1866), Estudios sobre una ley para el uso general del mar para la navegación y para los puertos, Barcelona: Establecimiento Tipográfico del Lloyds Español, 235 págs.

CUNNINGHAM, J. T., (1895), The inmature fish question, Journal of the Marine Biological Association, 3, págs. 54-77.

DUHAMEL DU MONCEAU, H.L., (1769-1782), Traité général des pêches et histoire des poissons, Tres tomos, Paris: Saillant & Nyon.

FERNÁNDEZ DÍAZ, R Y C. MARTÍNEZ SHAW, (1984), "La pesca en la España del Siglo XVIII. Una Aproximación Cuantitativa", Revista de Historia Económica, Año II, nº 3, págs. 183-201.

FERNÁNDEZ DÍAZ, R Y C. MARTÍNEZ SHAW, (1995), "Las Revistas de Inspección de la Matrícula de Mar en el siglo XVIII", en Martínez Shaw, C. (ed.), El Derecho del Mar en la España Moderna, Granada: Servicio de Publicaciones de la Universidad de Granada, págs. 241-271.

FERNÁNDEZ DURO, C., (1866a), Almadrabas. Reseña histórica de su empleo en las costas de España y reglamento para su régimen, Madrid: Estbto. Tip. de Estrada, Díaz y López, 108 págs.

FERNÁNDEZ DURO, C., (1866b), Estudios sobre la pesca con el arte denominado parejas del bou y reglamento para su régimen, Madrid: Estbto. Tip. de Estrada, Díaz y López, 102 págs.

FERNÁNDEZ DURO, C, (1868), Anuario de la Comisión Permanente de Pesca para 1868, Madrid: Tipografía de Estrada, 567 págs.

FERNÁNDEZ DURO, C, (1869), Anuario de la Comisión Permanente de Pesca para 1869, Madrid: Tipografía de Estrada, 541 págs.

FERRET, Z., (1819), Exposición histórica de las causas que más han influido en la decadencia de la Marina española e indicación de algunos medios para restaurarla, Barcelona: Roca y Gaspar, 176 págs.

FULTON, T.W., (1891), "The Scientific Work of the Fishery board for Scotland", Journal of the Marine Biological Association, 1, págs. 75-89.

GAMEZ DUARTE, F., (2006), El desafío insurgente. Análisis del corso hispanoamericano desde una perspectiva peninsular: 1812-1828. Logroño: Servicio de Publicaciones de la Universidad de La Rioja, 631 págs.

GARCÍA DEL HOYO, J.J., (2002), Liberalización Pesquera y Sobreexplotación en la Andalucía Atlántica de la primera mitad del siglo XIX, Huelva: Servicio de Publicaciones de la Universidad de Huelva, 61 págs.

GARCÍA DEL HOYO, J.J., (2006), "Economía Clasica, Liberalización Pesquera y Sobreexplotación en Andalucía", en I Conferencia Internacional de Historia de la Pesca en el Ámbito del Estrecho, Puerto de Santa María, 1-5 de junio de 2004, Sevilla: Consejería de Agricultura y Pesca de la Junta de Andalucía, págs. 961-1002.

GARCÍA SOLÁ, F., (1880), Memoria sobre la Industria y Legislación de Pesca que comprende desde el año 1874 al 1879, Madrid: Tipografía de Estrada, 840 págs.

GARCÍA SOLÁ, F., (1888), Idea General de la Pesca Marítima en España, Madrid: Imp. Vda. e hija de Gómez Fuentenebro, 26 págs.

GARSTANG, W., (1900), "The Impoverishment of the Sea. A Critical Summary of the Experimental and Statistical Evidence bearing upon the Alleged Depletion of the Trawling Grounds", Journal of the Marine Biological Association, 6, págs. 1-69.

GIRALDEZ RIVERO, J., (1991), "Fuentes estadísticas y producción pesquera en España (1880-1936): Una primera aproximación", Revista de Historia Económica, Año IX, nº 3, págs. 513-531.

GRACIA CÁRCAMO, J., (1984), "La evolución de la economía pesquera española en el siglo XVIII", Letras de Deusto, 30, 111-129.

GRAFE, R., (2003) 'The globalisation of codfish and wool: Spanish-English-North American triangular trade in the early modern period', LSE Economic History Working Paper Series No.72/03, 37 págs.

GUTIERREZ VELA, R., (1886), Memoria sobre la Industria y Legislación de Pesca que comprende desde el año 1879 al 1884, 2 vol., Madrid: Imprenta Vda. Fuentenebro, 903 págs.

GUTIÉRREZ VELA, R., (1894), "Estadística de Pesca. Año 1892", Revista de Pesca Marítima, Folleto separado, 56 págs.

HERRERA, R., (1887), "La Industria Pesquera en Ayamonte e Isla Cristina", Revista de Pesca Marítima, págs. 153-160 y 170-176.

IBÁÑEZ, J. M., (1844-1845), Tratado Elemental de Estadística, 2 tomos, Madrid: Imp. del Colegio de Sordomudos. Para las referencias se ha utilizado la edición realizada en 2006 por el Instituto Nacional de Estadística, Madrid: INE, 303 y 356 páginas.

LACOMBA, J.A., (2006), “El sector pesquero andaluz en el último cuarto del XIX: una fase de cambios y transformaciones. Una aproximación, Revista de Estudios Regionales, 75, págs. 129-150.

LON ROMEO, E., (1950), Trafalgar. Papeles de la Campaña de 1805, Zaragoza: Institución Fernando el Católico, 374 págs.

LÓPEZ CAPONT, F., (1997), La faceta pesquera del Padre Sarmiento y su época, Pontevedra: Caixa Pontevedra, 105 págs.

LÓPEZ LOSA, E., (2000), “La pesca en el País Vasco. Una visión a largo plazo (siglos XIX y XX)”, Itsas Memoria, Revista de Estudios Marítimos del País Vasco, 3, págs. 239-276.

LÓPEZ LOSA, E., (2005), “El Estado, la Marina y el Sector Pesquero en España durante los siglos XVIII y XIX”, en VIII Congreso de la Asociación Española de Historia Económica, 13-16 de septiembre de 2005, Santiago de Compostela.

LÓPEZ MIGUEL, O. Y M. MIRABET CUCALA, (1995), “La Institucionalización de la Matrícula de Mar: Textos Normativos y consecuencias para la gente de mar y maestranza”, en Martínez Shaw, C. (ed.), El Derecho del Mar en la España Moderna, Granada: Servicio de Publicaciones de la Universidad de Granada, págs. 217-239.

LLOVET, J., (1980), La Matrícula de mar i la província de marina de Mataró al segle XVIII, Mataró: R. Dalmau, 213 págs.

MARTÍN GARCÍA, A., (1999), “Entre el mar y la muerte. Procedencias, condiciones de vida

y mortalidad de los navegantes en el Real Servicio (1776-1804)”, Espacio, Tiempo y Forma, Serie IV, H.° Moderna, t. 12, págs. 415-441

MEREDIZ MORENO, A., (2004), Historia de la Estadística Oficial como Institución Pública en España, Sevilla: Instituto de Estadística de Andalucía, 323 págs.

MIRAVENT Y SOLER, J., (1850), Memoria sobre las pescas que se cultivan en las costas meridionales de España desde el Cabo de San Vicente hasta el Estrecho de Gibraltar, Huelva: Imp. Reyes y Moreno, 91 págs.

MOREAU DE JONNES, M.A., (1856), Statistique de l’Industrie de la France, París: Guillaumin Libreries, 380 págs.

OCAMPO SUÁREZ-VALDÉS, J., (1990), Campesinos y artesanos en la Asturias preindustrial (1750-1850), Gijón: Gran Distribuciones Gráficas 2000, S.L., 376 págs.

OCAMPO SUÁREZ-VALDÉS, J., (1992), “Traineras, vaporas y motoras: cambio técnico y especialización pesquera en Asturias, 1880-1930”, en VII Congreso de la Asociación de Historia Económica, Zaragoza.

O’DOGHERTY, A., (1952), “La Matrícula de Mar en el reinado de Carlos III”, Anuario de Estudios Americanos, vol. IX, pág. 347-370.

PENSADO, J.L. (ed.), (1992), De los Atunes y de sus Transmigraciones y Sobre el Modo de Aliviar las Miseria de los Pueblos, Salamanca: Universidad de Salamanca, 166 págs.

RAMÍREZ, F. Y A. NAVARRETE, (1905), Monografía de la Pesca Marítima en España. Madrid: Imprenta Alemana, 48 págs.

RICART, J., (1895), “Los pescadores y la pesca en España”, Revista de Pesca Marítima, XI, págs. 21-27. Había sido publicado anteriormente en la Revista de Navegación y Comercio en 1894.

RODRÍGUEZ GONZÁLEZ, A. R., (1999), La Armada Española. La Campaña del Pacífico 1862-1871, Madrid: Editorial Agualarga, 142 págs.

SALAS LARRAZABAL, F.J., (1870), Historia de la Matrícula de Mar y Examen de Varios Sistemas de reclutamiento marítimo, Madrid: Imprenta de Fortanet, 468 págs.

SALAS, J. Y F. GARCÍA SOLÁ, (1876), Memoria sobre la Industria y Legislación de Pesca que comprende desde el año 1870 al 1874, Madrid: Imprenta de Fortanet, 741 págs.

SAMPAIO, A.P.F. DE, (1841), Elementos de la Ciencias Estadística, Madrid: Imp. Boix, 76 págs.

SANZ SERRANO, A., (1956), Resumen Histórico de la Estadística en España, Madrid: Instituto Nacional de Estadística, 224 págs.

SAÑEZ REGUART, A., (1791-1795), "Diccionario histórico de los artes de la pesca nacional", Madrid: Imprenta de la Vda. de Ibarra. Se ha utilizado la edicción facsimil de 1988 del Ministerio de Agricultura, Pesca y Alimentación.

SMITH, T.D., (1994), Scaling Fisheries: The Science of Measuring the effects of fishing 1855-1955, Cambridge: Cambridge University Press, 392 pags.

UZTARIZ, G., (1724), Teórica y práctica de Comercio y de Marina, Madrid: Imprenta de Antonio Sanz. Se ha utilizado la edición facsimil de la 2ª edición de 1742 preparada por Ed. Aguilar en 1968.

VÁZQUEZ LIJÓ, J. M., (2005), "Supervisión y control de los gremios de mar por las autoridades de Marina: los cabos celadores y otras disposiciones de la ordenanza de matrículas de 1751", en VIII Congreso de la Asociación Española de Historia Económica, 13-16 de septiembre de 2005, Santiago de Compostela.

VIRUELA, J., (1995), "Expansión y crisis de la actividad pesquera valenciana en el siglo XIX", Investigaciones geográficas, 13, págs. 117-134.

Anexo I

Evolución de la actividad pesquera en el Departamento de Marina de Cádiz según los "Estados de Pesca". Serie Original y serie corregida.

AÑO	SERIES ORIGINALES				SERIES CORREGIDAS				
	Pescado Capturado		Embar-caciones	Pesca-dores	Pescado Capturado		Embar-caciones	Pesca-dores	
	Arrobas	Rs. vn.			Arrobas	Rs. vn.			
1828-1829	447106	5738908			599727	8870110			(1)
1830-1831	461623	4247171	716	5638	614244	7378373	716	5638	(1)
1831-1832	377484	4018952	655	4817	377484	4018952	655	4817	
1833-1834	375094	4817659	1204	7433	527715	7948861	602	3717	(1)
1845-1846	678085	7933719	1587	12615	679445	7933719	850	5717	(2)
1846-1847	774181	9403306	939	7243	803064	9403306	953	5795	(3)
1848-1849	839501	6984778	1672	9475	909831	8707209	1029	5834	
1850-1851	1042270	10687489	2560	14100	1062297	10687489	1125	6316	(4)
1851-1852	1266849	10455987	2453	13383	1266849	10455987	1159	6513	(5)
1852-1853	1060481	8074408	2211	12228	1060785	8074408	1124	6233	(4)
1856-1857	696776	8546220	1239	6140	696776	8546220	1239	6140	
1857-1858	1475455	16911435	1531	7970	1475455	16911435	1393	7392	(6)
1859-1860	1027182	13339146	1652	8474	1027182	13339146	1652	8474	
1860-1861	1412914	17987501	1765	7298	1412914	17987501	1765	7298	
1861-1862	1231965	17474617	1671	7820	1231965	17474617	1671	7820	
1862-1863	1614441	22964930	1835	7987	1614441	22964930	1835	7987	
1863-1864	1393626	23574530	1783	8969	1393626	23574530	1783	8969	
1864-1865	1241570	18178420	1640	7269	1241570	18178420	1640	7269	
1865-1866	2321532	20593950	1807	7808	2321532	20593950	1807	7808	
1866-1867	895453	14259350	2100	7800	895453	14259350	2100	7800	
1867-1868	1068617	15944332	1854	7659	1068617	15944332	1854	7659	
1873									
1878			2568	13169			2568	13169	
1883	1684500	38374348	2214	11155	1684500	38374348	2214	11155	
1889	2536560	83096762	2718	14255	2536560	83096762	2718	14255	
1892	2407187	56443704	3120	15735	2407187	56443704	3120	15735	

(1) Faltan datos de Canarias. En la serie corregida se han asignado los valores correspondientes del Estado de Pesca más cercano temporalmente.
(2) Errores de agregación en las embarcaciones y pescadores en Algeciras, Málaga y Canarias.
(3) Faltan datos de flota y empleo en Málaga. Se ha asignado el valor correspondiente más cercano temporalmente. Errores de agregación en Algeciras, Canarias y Huelva.
(4) Errores de agregación en las embarcaciones y pescadores en Málaga.
(5) Errores de agregación en las embarcaciones y pescadores en Málaga y Motril.
(6) Errores de agregación en las embarcaciones y pescadores en Almería.

Fuentes: Véase la Tabla V en el texto

Anexo II

Evolución de la actividad pesquera en el Departamento de Marina de Cartagena según los "Estados de Pesca". Serie Original y serie corregida.

AÑO	SERIES ORIGINALES				SERIES CORREGIDAS				
	Pescado Capturado		Embar-caciones	Pesca-dores	Pescado Capturado		Embar-caciones	Pesca-dores	
	Arrobas	Rs. vn.			Arrobas	Rs. vn.			
1828-1829	541061	6220070			541061	6220070			
1830-1831	507390	5401845	2549	10830	507390	5401845	1274	6202	(1)
1831-1832	446484	5566628	1422	7967	446484	5566628	1422	7967	
1833-1834	381141	5410057	2448	11889	381141	5410057	1224	6857	
1845-1846	498324	5976836	1400	7044	498324	5976836	1400	7044	
1846-1847	492739	5205797	1601	7724	492739	5205797	1489	7302	(2)
1848-1849	500017	6480460	1929	9664	500036	6480460	1532	7814	(3)
1850-1851	545021	6629515	2052	10299	553198	6629515	1459	7450	(4)
1851-1852	629925	7011166	2358	11387	628130	7011166	1566	7729	(5)
1852-1853	658769	8168123	1563	8106	658295	8168123	1563	8106	(2)
1856-1857	426800	7373843	1412	7435	446798	7373843	1512	7435	
1857-1858	505682	8941506	1978	8688	505682	8941506	1683	7650	(4)
1859-1860	489478	9910354	1685	7643	489478	9910354	1685	7643	
1860-1861	603710	10940063	1683	7345	603710	10910063	1683	7345	
1861-1862	569509	14943146	1610	7437	569509	14943146	1610	7437	
1862-1863	614959	9295190	1538	7458	614959	9295190	1538	7458	
1863-1864	592178	9870170	1823	8495	592178	9870170	1823	8495	
1864-1865	530387	11344560	2035	9372	530387	11344560	2035	9372	
1865-1866	491501	7401500	2883	12208	491501	7401500	2883	12208	
1866-1867	520234	7902680	2189	9258	520234	7902680	2189	9258	
1867-1868									
1873									
1878			4570	19966			4570	19966	
1883	735146	27888532	4343	16345	735146	606272	4343	16345	
1889	617416	38829749	3992	17174	617416	38829749	3992	17174	
1892	656427	19934692	4524	15598	656427	19934692	4524	15598	

(1) Errores de agregación en embarcaciones y pescadores en Alicante y Valencia.
(2) Errores de agregación en embarcaciones y pescadores en Mataró.
(3) Errores de agregación en embarcaciones y pescadores en Palamós y Valencia.
(4) Errores de agregación en embarcaciones y pescadores en Alicante, Barcelona, Ibiza y Menorca.
(5) Errores de agregación en embarcaciones y pescadores en Alicante, Barcelona, Ibiza, Mallorca, Mataró y Menorca.
(6) Errores de agregación en embarcaciones y pescadores en Ibiza, Tarragona y Tortosa.

Fuentes: Véase la Tabla V en el texto

Anexo III

Evolución de la actividad pesquera en el Departamento de Marina de El Ferrol según los "Estados de Pesca". Serie Original y serie corregida.

AÑO	SERIES ORIGINALES				SERIES CORREGIDAS				
	Pescado Capturado		Embar-caciones	Pesca-dores	Pescado Capturado		Embar-caciones	Pesca-dores	
	Arrobas	Rs. vn.			Arrobas	Rs. vn.			
1828-1829	568862	4094106			777335	6056567			(1)
1830-1831	3855406	27588945	5949	35847	2583888	17971345	4684	25672	(2)
1831-1832	992919	5964889	3430	13408	992919	5964889	4430	23408	
1833-1834	558791	3961718	4866	21501	868509	5299511	4866	21501	(1)
1845-1846	1008509	4799511	4018	17988	1193005	6554990	2889	14692	(1)
1846-1847	827519	6811008	3754	17378	827519	6811008	3578	16259	(3)
1848-1849	571939	4807302	5000	22363	572183	4807302	3911	13094	(3)
1850-1851	1278825	8551778	3423	13798	1278825	8551778	3205	12290	(4)
1851-1852	1300626	7430017	3674	14020	1300626	7430017	3314	12488	(5)
1852-1853	1180261	7541033	4254	22318	1180261	7541033	3655	12618	(5)
1856-1857	3424656	22314860	4594	18710	3424656	22314860	3848	13874	(4)
1857-1858	4808350	37930582	4714	16578	4810350	37930582	4308	14955	
1859-1860	4337209	35626762	5270	18140	4337209	35626762	4810	15860	
1860-1861	4571277	49287403	4744	18284	4571277	48287403	4271	16008	
1861-1862	3931813	32904658	4373	16676	3931813	32904658	4373	16676	
1862-1863	3809504	38058370	4782	18578	3809504	38058370	4782	18578	
1863-1864	4861977	28207590	5095	18823	4861977	28207590	5095	18823	
1864-1865	3543382	46716790	5477	19502	3543382	46716790	5477	19502	
1865-1866	3067015	38606420	5658	19424	3067015	38606420	5658	19424	
1866-1867	2797286	29304790	5927	20500	2797286	29304790	5927	20500	
1867-1868									
1873	2955156	26893556	6153	20150	2955156	26893556	6153	20150	
1878			6879	33107			6879	33107	
1883	3458966	79262820	9178	38710	3458966	79262820	9178	38710	
1889	3681221	104604479	8015	32481	3681221	104604479	8015	32481	
1892	4121949	76585976	7092	35864	4121949	76585976	7092	35864	

(1) Faltan los datos de Bilbao y San Sebastián.
(2) Faltan los datos de Bilbao. Errores de agregación en embarcaciones y pescadores en San Sebastián.
(3) Errores de agregación en embarcaciones y pescadores en San Sebastián.
(4) Faltan los datos de pescadores en Bilbao. Errores de agregación en embarcaciones y pescadores en San Sebastián.
(5) Faltan los datos de pescadores y valor de las capturas en Bilbao. Errores de agregación en embarcaciones y pescadores en San Sebastián.

Fuentes: Véase la Tabla V en el texto.

Anexo IV

Evolución de la actividad pesquera en el conjunto del Estado según los "Estados de Pesca". Serie Original y serie corregida.

AÑO	SERIES ORIGINALES				SERIES CORREGIDAS			
	Pescado Capturado		Embar-caciones	Pesca-dores	Pescado Capturado		Embar-caciones	Pesca-dores
	Arrobas	Rs. vn.			Arrobas	Rs. vn.		
1828-1829	1557029	16053084			2126596	23109207		
1830-1831	4824419	37237961	9214	52315	3705522	30751563	6674	37512
1831-1832	1816887	15550469	5507	26192	1816887	15550469	6507	36192
1833-1834	1315026	14189434	8518	40823	1777365	18658429	6692	32075
1845-1846	2184918	18710066	7005	37647	2370774	20465545	5139	27453
1846-1847	2094439	21420111	6294	32345	2123322	21420111	6020	29356
1848-1849	1911457	18272540	8601	41502	1982050	19994971	6472	26742
1850-1851	2866116	25868782	8035	38197	2894320	25868782	5789	26056
1851-1852	3197400	24897170	8485	38790	3195605	24897170	6039	26730
1852-1853	2899511	23783564	8028	42652	2899341	23783564	6342	26957
1856-1857	4548232	38234923	7245	32285	4568230	38234923	6599	27449
1857-1858	6789487	63783523	8223	33236	6791487	63783523	7384	29997
1859-1860	5853869	58876262	8607	34257	5853869	58876262	8147	31977
1860-1861	6587901	78214967	8192	32927	6587901	77184967	7719	30651
1861-1862	5733287	65322421	7654	31933	5733287	65322421	7654	31933
1862-1863	6038904	70318490	8155	34023	6038904	70318490	8155	34023
1863-1864	6847781	61652290	8701	36287	6847781	61652290	8701	36287
1864-1865	5315339	76239770	9152	36143	5315339	76239770	9152	36143
1865-1866	5880048	66601870	10348	39440	5880048	66601870	10348	39440
1866-1867	4212973	51466820	10216	37558	4212973	51466820	10216	37558
1867-1868								
1873								
1878			14017	66242			14017	66242
1883	5878612	145525700	15735	66210	5878612	118243440	15735	66210
1889	6835198	226530990	14725	63910	6835198	226530990	14725	63910
1892	7185563	152964372	14736	67197	7185563	152964372	14736	67197

Fuentes: Véase la Tabla V en el texto

Capítulo 18

El estudio de la mortalidad en España a principios del siglo XX

SONIA DE PAZ COBO
JUAN MANUEL LÓPEZ ZAFRA
Universidad Complutense de Madrid

Antecedentes españoles en el estudio de la mortalidad

Tal y como señalamos en de Paz Cobo y López Zafra (2006), el estudio sistemático, científico de la mortalidad en España, y de sus causas y efectos sobre la población, no comienza sino a partir de la obra de Merino (1866) (y, asimismo, de cierta aportación de Coll que señala Sorribas en su obra de 1882, pero de la que desgraciadamente ya entonces no quedaba constancia alguna); como es sabido, esta autor se basa en la *Memoria sobre el movimiento de la población de España en los años 1858, 1859, 1860 y 1861* de la Junta de Estadística, cuyo principal problema metodológico estribaba en la agrupación de los fallecidos en bloques de cinco años. Más de 20 años antes, la Compañía General Española de Seguros había publicado una Tabla de Seguros sobre la Vida, GES (1841), pero no se recoge en ellas otra experiencia que la extranjera, reflejo del término medio de las publicadas en diversos lugares de Europa; cierto es que entre las tablas que cierran la publicación ninguna de ellas es de mortalidad.

Posteriormente, Sorribas (1863) publica la que puede considerarse como la obra española más importante en este ámbito en el s. XIX, al apoyarse no sólo en datos oficiales sino asimismo en lo relativos a tres poblaciones de cierta importancia para la fecha. Vuelve a enfrentarse al problema de la agrupación quinquenal de las defunciones, que resuelve empleando ciertos procedimientos actuariales y que ajusta a los datos primarios que obtiene de las tres poblaciones reseñadas.

Podemos afirmar, sin temor a equivocarnos, que estamos ante dos trabajos muy serios, de alto rigor científico, y que recogen plenamente las aportaciones científicas que en ese momento existían en Europa; sin embargo, el marco actuarial del estudio de la mortalidad experimentaría un cambio sustancial con la obra de Makeham (1859), quien añade a la conocida desde 1825 ley de Gompertz una constante para reflejar que el tanto instantáneo de mortalidad depende no sólo de la edad. Y es precisamente la obra de los hermanos Puyol Lalaguna de 1911 la primera en España que recoge tal aportación.

La *Tabla de mortalidad española ajustada analíticamente* de los hermanos Puyol Lalaguna de 1911

Comienzan los autores con un lamento que coincide con el expresado por D. Miguel Merino casi 50 años antes, en sus *Reflexiones y conjeturas sobre la ley de mortalidad en España* de 1866, en lo relativo a la inexistencia de tablas de mortalidad españolas "basadas sobre la experiencia de las Compañías de seguros". Se quejan asimismo de las grandes dificultades que "la falta de datos, y aún más, la inexactitud casi inevitable de los suministrados por los censos" provocan en la elaboración de una tabla correcta, junto con las deficiencias que surgen de una población en la que la emigración es una característica fundamental.

Es necesario hacer aquí una breve reseña de los autores de este trabajo. Mateo fue un destacado actuario que ya en 1909 había publicado su *Manual de seguros sobre la vida*, una obra de carácter elemental que permitió la divulgación de de los principio básicos del seguro; desde 1912 y hasta 1922 fue Jefe Técnico de la Comisaría General de Seguros, dando forma durante ese período a la fundamental acción tutelar del Estado sobre el sector asegurador; su interés por la enseñanza de la profesión queda de manifiesto cuando recibe el encargo, en 1914, por parte del Instituto de Actuarios (colegio profesional que regula el acceso a la profesión desde hace casi 100 años), de analizar la organización de los estudios actuariales en el extranjero, labor que culmina con una propuesta de lo que podría ser la enseñanza del Seguro en España. La importancia de D. Mateo en al ámbito actuarial queda reflejada en el homenaje que el Instituto de Actuarios, a través de su revista Anales, le brindó en su número correspondiente a los años IV y V, publicado en 1947.

Su hermano José María había publicado también en 1910 la obra *Las casas baratas*, cuyo subtítulo rezaba "*Proyectos, planos y presupuestos. La amortización. El seguro. Las sociedades constructoras y el Instituto Nacional de Previsión.*" Editada por la Oficina de Trabajo de la "Acción Social Popular" de Barcelona, recogía los aspectos fundamentales (entre ellos, el del seguro) relativos a las denominadas "casas baratas", antecedentes de las VPO. La primera Ley que en España recogía las características de este tipo de viviendas se promulgó en 1911 (aunque el primer proyecto fue aprobado en el seno del Instituto de Reformas Sociales en 1907).

En cuanto a la metodología y los datos empleados, los autores suponen que la ley de supervivencia española coincide con la de Makeham (con la salvedad, según ellos, de que la mortalidad de la población general no puede ser la misma que la de la población de asegurados) que ajustan a los datos obtenidos de los censos generales por edades de 1877 y 1887 del Institutos Geográfico y Estadístico; la razón de no emplear el último de ellos, más reciente, es a causa de reflejar variaciones debidas al cólera y a la gripe que asolaron España esos años. En cuanto a la estadística de defunciones, surge el sempiterno problema del agrupamiento por tramos quinquenales, por lo que los autores optan por ajustar proporcionalmente a la tasa de mortalidad de cada edad. Otro problema en los datos era la enorme diferencia de las edades terminadas en 9, 0 y 1 en relación a las inmediatamente anteriores y posteriores, por lo que optaron por proceder con ellas como con anteriormente expuesta acerca de los fallecidos. Tras observar que efectivamente la fórmula de Makeham

($l_x = ks^x g^{c^x}$) ajusta correctamente la supervivencia a cada edad para los adultos, y no así en los primeros años de la vida (pág. 7), los autores pasan a ajustar las constantes s, g y c de la referida, que

> *"dependerán del clima, salubridad, higiene, sexo, etc, de todo lo que influye en la prolongación o acabamiento de la vida humana."*

Tal y como señalan en la pág. 8, los autores siguen el método de las sumas de King y Hardy publicado en el vol. 22 (1879-1881) del *Journal of the Institute of Actuaries* (pp. 191-231), empleando como edad límite o superior los 67 años, de acuerdo con lo establecido por de Moivre.

Así, estiman finalmente los valores de las tres constantes previamente señaladas en la forma siguiente:

s = 0,9911431; g = 0,9990522; c = 1,0977401.

Del ajuste efectuado se comparan las cifras de fallecimientos previstas con las realmente observadas, obteniéndose una subestimación agregada de 4597 individuos, lo que supone una exactitud del 91 por ciento. Por último (pág. 13), tras determinar el número de supervivientes a la edad de veinte años (que depende del número arbitrario tomado como cifra inicial, que los autores fijaron en 600.000 individuos), deducen la última constante del modelo, k, como 399114. Queda pues la fórmula de supervivencia a la edad x ajustada analíticamente por los hermanos Puyol Lalaguna establecida definitivamente como

$$l_x = 399114 \times 0.9911431^x \times 0.9990522^{1.0977401^x}.$$

Finaliza este importante trabajo con las últimas páginas (14 a 21) dedicadas por los autores a exponer su Tabla de mortalidad española según el censo de 1877, así como dos cuadros en los que comparan los valores de las tasas anuales de mortalidad, por un lado, y de las rentas vitalicias de una peseta (calculadas a un tipo del 3,5%), por otro, por ellos obtenidos con los de los que se siguen de algunas de las principales tablas empleadas en operaciones de seguros, así como con un explícito diagrama con las curvas que permiten visualizar los resultados.

Fuentes Martiáñez y las *Tablas de mortalidad, supervivencia, vida media y vida probable* de 1927. El comienzo de una nueva etapa.

Es en 1927 cuando D. Mariano Fuentes Martiáñez, jefe del Cuerpo facultativo de Estadística, publica la citada obra. Puede considerarse, sin lugar a dudas, que es ésta la que cierra la andadura inicial del análisis de la supervivencia español y que con ella la estadística oficial española se integra de lleno, definitivamente, en la corriente científica europea en este ámbito, equiparándose en rigor, calidad y conocimientos a las de los países más avanzados de nuestro entorno, como pudieran ser el Reino Unido, Francia, Bélgica o Suecia.

Ya en el prólogo señala el autor que "nos sugirió la idea de llegar a la formación de unas verdaderas tablas de mortalidad. Tales cálculos nunca se han intentado seriamente en nuestra patria (…)", efectuando una no velada crítica a los intentos anteriores de Merino, Sorribas y sobre todo Puyol, cuando éste, tal y como hemos señalado, sobrepasa con mucho los trabajos de los anteriores. La notación empleada por el autor es ya completamente estándar y lógicamente coincide con la internacionalmente aceptada, y la reseñamos en el siguiente cuadro 1.

Define la mortalidad (pág. 16) como una función tanto de las defunciones como de la población, para justificar de ese modo el empleo del período 1908-1923 y determinar el promedio anual de fallecidos para cada edad. Para hallar tal promedio prescinde del año 1918, debido a la incidencia de la gripe en la mortandad. En cuanto al dato de población, emplea los censos de 1910 y 1920, justificando a continuación el período de observación (1908-1923) en el criterio de Fan, pues según el autor deben ser considerados "los 10 años del período intercensal y 3 anteriores y 3 posteriores a ese mismo período".

Símbolo	Explicación
l_x	Supervivientes a la edad x
d_x	Fallecidos a la edad x
P_x	Población censada a la edad x
p_x	Probabilidad de sobrevivir a la edad x
q_x	Probabilidad de morir a la edad x
p^t_x	Probabilidad de un individuo de edad x de vivir t años
q^t_x	Probabilidad de un individuo de x años de morir antes de cumplir x+t años
μ_x	Tasa instantánea o fuerza de mortalidad
L_x	Semisuma de l_x y l_{x+1}
K, s, g, c	Constantes de la fórmula de Makeham
e^0_x	Vida media o esperanza de completa de vida a la edad x
$V^{(x)}_p$	Vida probable a la edad x

Cuadro 1. Símbolos de conmutación empleados por Fuentes (1927)

Por vez primera aparece en España la moderna concepción de la tabla de mortalidad (pág. 17), que "debe servir al mismo tiempo de tabla de supervivencia y en ella han de hacerse constar los datos, no menos interesantes, de la vida media y la vida probable en las distintas edades del hombre", más allá de la "mera exposición por edades de los correspondientes coeficientes de mortalidad."

Entra a continuación el autor a señalar la formulación empleada en el estudio, con múltiples ejemplos muy didácticos sobre los ajustes que deben efectuarse para conseguir cifras adecuadas. Así, por ejemplo, define de forma directa

$$q_x = d_x / P_x ,$$

lo aplica al censo de 1920 para los individuos de 25 años

$$q_{25} = 3043/357735 = 0.0085 ,$$

y a continuación explica el concepto que hoy conocemos como edad actuarial para corregir y conseguir

$$q_x = d_x / \left(P_x + d_x/2\right) \quad ; q_{25} = 3043 / \left(357735 + 3043/2\right) = 0,00847 .$$

Es en la página 20 cuando plantea el procedimiento de Knapp y Zenner para determinar la probabilidad exacta de sobrevivir a la edad x, lo aplica a x=25 años, y determina a continuación por complemento el coeficiente de mortalidad, mucho más exacto que el anterior. Sin embargo, tal y como señala el propio autor, en la práctica este procedimiento puede ser inaplicable por defectos en el acta o registro de defunción, pues exige conocer no sólo la edad del finado sino también con total precisión su fecha de nacimiento, la cual no constaba por entonces en la mayoría de las ocasiones.

Pasa entonces revista a diversos métodos de ajuste y cálculo de la mortalidad (seguimiento durante 100 años de 100.000 nacidos, cálculos bajo hipótesis de estacionariedad de la población, método de Halley, de Hermann), señalando las ventajas e inconvenientes de cada uno de ellos (con ejemplificación numérica) para justificar finalmente (pág. 27) el empleo del de Quetelet, del diezmo mortuorio o de la probabilidad de muerte ($q_x = d_x / \left(P_x + d_x/2\right)$ anteriormente citado). Entra entonces en la explicación de los problemas derivados de tal recargo de ciertas edades, pues "la simpatía por las cifras redondeadas introduce siempre un elemento pertubador, que es preciso corregir" (pág. 28), y en su solución mediante los distintos métodos de ajuste, de los que ofrece una detallada explicación de los diversos empleados por entonces (métodos gráficos, de Highman, de Finlaison, de interpolación parabólica, de Altenburger, de Woolhouse, de Witstein y de la fórmula de Makham); decide emplear el autor los de interpolación parabólica, por su sencillez, y el de Makeham. Efectúa entonces pruebas de los diversos métodos de ajuste a lo largo de las páginas 30 a 39.

Más adelante se introduce en los conceptos de vida media "o esperanza matemática de la vida", mediante

$$e_0^x = \frac{1}{2} + \left(l_{x+1} + l_{x+2} + \cdots + l_{x+n}\right) / l_x ,$$

aplicándolo a modo de ejemplo (pág. 40) a los casos

$$e_{15}^0 = 3118037/68712 = 45,88 \text{ años},$$

y

$$e_{50}^0 = 1001440/51203 = 20,06 \text{ años},$$

y en el de vida probable, "el tiempo t que ha de transcurrir para que un cierto número de personas de edad determinada se reduzca a la mitad",

$$n = l_{x+r} - \frac{1}{2} l_x ; \ n' = \frac{1}{2} l_x - l_{x+r+1}$$

$$n + n' = l_{x+r} - l_{x+r+1} = d_{x+r}$$

señalando el autor que "si en el intervalo de 1 año, que media entre x+r y x+r+1, se han producido d_{x+r} fallecimientos, se puede admitir que un fallecimiento se producirá en el lapso de tiempo $1/d_{x+r}$, luego los n fallecimientos de diferencia que existen entre l_{x+r} y $\frac{1}{2}l_x$ tardarán en producirse n/d_{x+r}, que será la fracción de año que habremos de agregar a r para obtener la vida probable a la edad x. Es decir que

$$V_p^{(x)} = r + \frac{n}{d_{x+r}},$$

que es la fórmula empleada en nuestros cálculos.” (pág. 41)

Introduce asimismo el concepto de vida normal, edad de muerte más frecuente de los individuos de una población en un momento dado. Entra a continuación (pág. 44 y ss) en la comparación de las series primitiva y ajustada, para observar las diferencias que por exceso o por defecto se dan, para cada edad, en cuanto al número de supervivientes de una población teórica de 100.000 individuos en su origen. Se observa en general que la serie ajustada presenta mayores valores (más supervivientes) que la primitiva, por lo que se plantea la necesidad de “buscar un ajustamiento más perfecto entre ambas series”, cosa que acometerá, mediante el empleo de las aportaciones de Makeham, en la segunda parte del trabajo.

Tras plantear diversas expresiones de la vida media empieza la segunda parte de la obra (pág. 50 y ss) dedicada a la fórmula de Makeham.

Efectúa el autor una introducción histórica al estudio moderno de la función de mortalidad, iniciada con de Moivre (“el número de personas existentes en cada edad es proporcional al número de años de diferencia entre la edad límite de una tabla de mortalidad y la edad considerada”), seguida con Gompertz y jalonada (entonces) con Makeham. Fija el concepto de fuerza o tasa instantánea de mortalidad (pág. 54) como

$$\mu_x = -\frac{d \ln l_x}{d_x},$$

del que ofrece una muy prolija explicación a lo largo de 5 páginas, y entra a continuación en la descomposición de la fórmula de Gompertz, para alcanzar finalmente (pág. 63) al de Makeham, de la que ofrece una explicación de unas 4 páginas. A continuación se entretiene en ofrecer una aplicación de la misma, determinando en primer lugar las constantes a través de las fórmulas de Milne (pág. 70 a 90) y de King y Hardy (pág. 90 a 106), en segundo lugar los coeficientes de superviviencia y de mortalidad (pág. 107 y 108), y en tercer y último lugar la tasa instantánea de mortalidad (pág. 108 a 120, por bloques de edad: de 23 a 75 años, de 10 a 22, de 0 a 9, y de 76 a 102). De este modo, y tras un análisis rigurosísimo y unas explicaciones de gran poder didáctico, alcanza por fin la ansiada tabla de mortalidad en la página 120 de la obra, bajo la denominación de Estado nº 4. Sin embargo, antes de dar por finalizado el trabajo, el autor (pág. 127) efectúa un conjunto de consideraciones finales. En ellas afirma la bondad del Estado nº 4 o tabla de mortalidad definitiva, señalando sin embargo la publicación complementaria (Estado nº 3) de la tabla de mortalidad obtenida por interpolación parabólica, “para que sus cifras sean confrontadas con la experiencia de las compañías de seguros domiciliadas en España” (pág. 127).

Añade la posibilidad de que “seguramente” sus estimaciones acerca de la mortalidad excedan la de las citadas compañías aseguradoras, aduciendo que “la población asegurada es más sana que la población total del país”; este tipo de consideración favorable hacia la selección positiva era relativamente normal en la época, al considerarse que los más proclives a contratar un seguro de vida (o de cualquier otra índole) eran los más pudientes, y en ese sentido posiblemente con mayores expectativas de vida que el resto de la población.

Defiende el autor sus tablas también con un argumento de futuro, al indicar que las tablas calculadas sobre la población total puedan ser de gran utilidad a los seguros de carácter social (pensiones, vejez, enfermedad, invalidez), en los que en ocasiones es el Estado el único gestor, como es el caso de las clases "de recursos económicos más precarios" (pág. 130).

Muestra por último unas comparaciones internacionales en cuanto a la vida media, aunque lo más interesante son las consecuencias que extrae en la relación de los coeficientes de mortalidad por él obtenidos y los que manejan las entidades aseguradoras tanto españolas como extranjeras. Así, por ejemplo, las tablas francesas (AF y RF) presentan una menor mortalidad en los diez primeros años de vida; mayor entre los 11 y los 17 años; desde los 18, la RF mantiene valores siempre inferiores, mientras que la AF presenta coeficientes superiores en el tramo de 38 a 67 años, siendo menores de nuevo desde los 68 años en adelante. Del mismo modo, la tabla inglesa de las 17 compañías (que no considera edades por debajo delos 10 años), desde los 11 hasta los 64 presenta coeficientes de mortalidad mayores a las de Fuentes, de forma similar a lo que ocurre (aunque sólo hasta los 45) en la *American Experience Table of Mortality*, muy empleada en los EE.UU.

Conclusiones.

Culmina de este modo ejemplar un periplo que comenzaba más de 60 años atrás gracias a Coll y a Merino, proseguía con Sorribas y se asentaba con los hermanos Puyol Lalaguna. El estudio de la mortalidad española sigue desde ese momento la senda marcada por los países europeos con mayor tradición actuarial, y conforma un preludio al análisis sistemático que llevará a cabo el Instituto Nacional de Estadística a partir de la finalización de la Guerra Civil.

Bibliografía

COMPAÑÍA GENERAL ESPAÑOLA DE SEGUROS CONTRA INCENDIO (1841) *Tablas de seguros sobre la vida*. Imp. de don Eusebio Aguado, Madrid.

FUENTES MARTIÁÑEZ, M. (1927) *Tablas de mortalidad, supervivencia, vida media y vida probable.* Madrid.

DE PAZ COBO, S; LÓPEZ ZAFRA, JM (2006) "El estudio de la mortalidad en España en el Siglo XIX" *Historia de la Probabilidad y la Estadística*, Vol. III. Delta, Madrid.

MAKEHAM, W. (1859). "On the Law of mortality and the construction of annuity tables". Journal of the Institute of Actuaries and Assurance Magazine, 8.

MERINO, MIGUEL (1866) Reflexiones y conjeturas sobre la ley de mortalidad en España. Eduardo Cuesta, Madrid.

PUYOL LALAGUNA, M.; PUYOL LALAGUNA, JM. (1911) Tabla de mortalidad española ajustada analíticamente. Ricardo F. de Rojas, Madrid.

SORRIBAS Y ZAIDÍN, JA (1883). Memoria dilucidando un tema de seguros sobre la vida. Jaime Jepús, Barcelona.

Capítulo 19

Historia de las tablas de mortalidad españolas y su evolución

JUAN ESCUDER BUENO,
ROBERTO ESCUDER VALLÉS
Universidad de Valencia
ANGEL VEGAS MONTANER
Universidad de Alcalá

1.- Breves referencias sobre tablas de mortalidad

Aunque son sobradamente conocidas en el ámbito actuarial, consideramos conveniente hacer una referencia previa al concepto de tabla de mortalidad, por supuesto sin propósito de exhaustividad, sino simplemente de situar el ámbito en el que se va a desarrollar este trabajo.

Una de las características básicas del seguro de vida es la fuerte interconexión entre la probabilidad de fallecimiento de una persona y su edad, relación que, en grandes carteras, ha conducido a la elevada estabilidad y solvencia que caracteriza a las compañías de seguros de vida. Tal dependencia se puede modelizar de distintas formas, y se suele configurar en forma de tabla discreta denominada tabla de mortalidad.

Toda tabla de mortalidad incluye como parámetro inicial l_x, que representa el número de supervivientes a la edad $\boldsymbol{x}$. Elementos básicos a partir de él son: d_x**,** que representa el número de fallecidos a la edad x, q_x que es la probabilidad de fallecer a la edad x, p_x o probabilidad de que un individuo de edad $\boldsymbol{x}$ sobreviva al cabo de un año, e_x que representa la esperanza de vida de una persona de edad x**,** y μ_x que constituye el tanto instantáneo de mortalidad, parámetro que en la mayoría de los modelos paramétricos modeliza la ley de supervivencia mediante integración de la correspondiente ecuación diferencial.

Además de la edad, existe un factor de riesgo menos relevante pero importante que es el sexo del expuesto al riesgo. Por ello, se elaboran tablas de mortalidad para hombres y mujeres diferenciadas. En los momentos actuales, de fuerte dependencia de lo políticamente correcto, se ha pretendido por parte de instancias oficiales de orden político (como la Comisión Europea, por ejemplo) impedir el tratamiento diferenciado de hombres y mujeres en la aplicación de las tablas de mortalidad (véase Sáez de Jáuregui (2007)). Pero técnicamente es indiscutible la existencia de diferencias, y cada vez más evidentes una vez que se va prolongando la esperanza de vida como consecuencia de las mejoras sanitarias y asistenciales.

Otro elemento diferenciador es la generación dando lugar a las llamadas tablas generacionales. También pueden elaborarse tablas diferenciadas de muerte y supervivencia, según sea el riesgo asegurado, tablas que incorporan un recargo de seguridad que privilegia el supuesto contrario al asegurado. Además diferentes criterios de graduación (McCutcheon, J.J. (1987)) así como diferentes métodos generales en cuanto a su construcción, como detalla minuciosamente el profesor F. Insolera (1950) y también el profesor A. Vegas (1981).

En forma de anexo matemático se presentan los modelos fundamentales que se han utilizado para el ajuste o graduación de las tablas de mortalidad.

2.- Antecedentes históricos

2.1.- Desde la antigüedad hasta el siglo XVII

Las tablas de mortalidad tal y como hoy las conocemos son un evento relativamente moderno, no podemos hablar de "verdaderas tablas de mortalidad" por lo menos hasta los siglos XVII y XVIII, aunque no debemos olvidar que la preocupación por la mortalidad siempre influyó en la vida y pensamiento de la humanidad. Tampoco es nuestra intención estudiar esta época.[*]

No existen noticias fidedignas escritas sobre lo que en la antigüedad se realizaba en cuanto a registros de la mortalidad, debido fundamentalmente a la destrucción de las bibliotecas de Efeso y Alejandría. Por ello no podemos conocer el tratamiento que sobre el particular realizaban asirios, hebreos, egipcios y fenicios e incluso griegos y romanos (hasta épocas ya avanzadas).

Sin embargo, si que existen hitos, noticias y relatos sobre la actividad aseguradora entendida como una prevención frente a los riesgos y consecuencias de la mortalidad, hambrunas, tempestades etc.

Como noticias esporádicas, recogida de relatos, podemos entresacar hasta la Baja Edad Media, que:

- Como referencia más antigua, en el antiguo Egipto de las pirámides de Gizeh existía una mutualidad de riesgos mutuos organizada por los maestros artesanos que elaboraban las esculturas y elementos ornamentales utilizados en dichas edificaciones.

- En Babilonia existía una asociación mutual de comerciantes que sufragaba los gastos ocasionados a las embarcaciones por tempestades y accidentes, y que los hebreos compensaban las pérdidas de animales de sus individuos con otros de similares características.

- Los griegos también practicaban el seguro. Parece ser que llevaban estadísticas de los que fallecían luchando para poder socorrer a viudas y huérfanos.

[*] La información correspondiente a dicho período, excepto algunas referencias concretas, la hemos entresacado fundamentalmente de la enciclopedia actuarial Teugels-Sundt Editors (2004), y de trabajos "on line" de paginas web.

- En cuanto a los primeros intentos de controlar los movimientos naturales de población (nacimientos y defunciones) de que tenemos noticia, éstos datan de la Roma del año 578 aC y de la China del siglo III aC, aunque tenían motivaciones militares e impositivas.

- En Roma tuvo especial importancia el censo ordenado por Augusto alrededor del año cero de nuestra era, por su relación con el nacimiento de Cristo. Los censos practicados fueron utilizados con motivos impositivos, sobre todo a partir de la época de Servio Tulio. Con anterioridad al nacimiento de Cristo, el año 40 anterior a nuestra era, se elaboró la conocida como tabla Macer en base a la Ley Faldiciana, del tribuno Faldicius. Sin embargo, es la tabla Ulpiana, elaborada en 220 de nuestra era, por el eminente jurista y prefecto pretoriano Domitius Ulpianus, el primer hito histórico reseñable con perfección técnica (como han demostrado los actuarios Mays para edades superiores a 25 años, Splicing para edades inferiores a 25, etc.).

- Mas tarde aparecieron las asociaciones mutuales del ejército romano y poco a poco se fueron extendiendo a otros estamentos, como para compensar a viudas y huérfanos, que se extendieron a las asociaciones gremiales incipientes de la Baja Edad Media, y también con la pretensión de ayudar a sus miembros en caso de enfermedad, incendio o viaje.

Centrándonos en España, en esta época hay diferencias sustanciales con Europa debidas a la dominación árabe (así tenemos constancia de los censos ordenados por Alhacam II, a finales del siglo X, y por Abd-el-Mumén, en el XI). Pero similarmente a lo acontecido en Europa también en la España cristiana del siglo XII (sobre todo en la Corona de Aragón) surgieron las denominadas Cofradías Gremiales que auxiliaban a sus miembros ante la enfermedad y con pensiones por vejez o invalidez. Los juros y censos, por el análisis de las cartas en que se otorgaban, contienen los principios fundamentales de las rentas vitalicias perpetuas o temporales. (Luis Benítez de Lugo (1955).

Otros acontecimientos a destacar con anterioridad al s. XVII, fueron:

- el censo ordenado por los Reyes Católicos a Alonso de Quintanilla en 1482,
- el realizado bajo mandato de Carlos V en 1541 sobre los vecinos de las 18 provincias de su reinado, y
- los efectuados en época de Felipe II, quién, además de los dos empadronamientos generales de 1587 y 1594, encargó en 1574 a Ambrosio de Morales la Descripción de los pueblos de España, cuyo trabajo todavía se conserva parcialmente en la Real Biblioteca de San Lorenzo del Escorial.

2.2.- Desde el siglo XVII hasta la actualidad

La elaboración de tablas de mortalidad como más o menos las conocemos hoy en día, podemos decir que comienza entre finales del siglo XVII y principios del XVIII con una metodología incipiente, siendo sobre todo a partir de la segunda mitad del siglo XIX cuando se consolida la metodología y podemos hablar propiamente de tablas de mortalidad.

Como hemos indicado, hasta el siglo XVII no existen documentos sobre los que podamos afirmar la existencia de verdaderas tablas de mortalidad, aunque si podríamos citar determinados tipos de datos censales publicados en Francia por Bodin en 1557, unas tablas de distribución de sexos en Inglaterra en 1592, y un censo realizado en Suecia en 1608.

Los grandes matemáticos, geómetras y astrónomos del siglo XVII se sintieron atraídos por el fenómeno biométrico. Así, De Witt, Hudde, Huygens, Halley, etc. intervinieron decisivamente en el desarrollo de la biometría actuarial. Incluso el gran Isaac Newton aparece en un acta dando su "imprimatur" a una tabla de mortalidad que presentaba los cálculos de rentas para el último superviviente sobre varias cabezas.

Se suele considerar como primera tabla "científica" la elaborada en 1662 por John Graunt, basada en experiencia de mortalidad inglesa de los periodos 1629-36 y 1647-60. Se trataba de una tabla simplificada, que sólo recogía las edades acabadas en 6 hasta la de 76 años. En base a esta tabla, Lodewijk Huygens (hermano de Christiaan) mostró, por primera vez, en 1669, cómo calcular la esperanza de vida.

El primer intento de calcular probabilidades sobre la vida humana lo llevaron a cabo en Holanda de Witt y Hudde, quienes en 1671 publicaron un trabajo sobre probabilidades y juegos de azar siguiendo los trabajos de Cardano, Christiaan Huygens, Pascal y Fermat, llegando a calcular el valor actual actuarial de una renta vitalicia (con errores).

El siguiente eslabón a reseñar sería la tabla de mortalidad Breslau, elaborada por el matemático, actuario y astrónomo Edmund Halley ("An estimate of the Degrees of the Mortality of Mankind"), publicada en 1693, en base a la experiencia de mortalidad del periodo 1687-91 producida en la ciudad de Breslau. Leibniz remitió tal estadística al conocido astrónomo Halley como secretario que era de la Royal Society of London. En base a dicho censo, Halley elaboró una tabla de mortalidad en el sentido moderno del término y calculó el valor actual actuarial de una renta vitalicia si bien con un algoritmo que se considera peor que el de Witt. La tabla Breslau, sobre la que Colin MacLaurin actuó como consultor, jugó un papel fundamental en la fundación de uno de los primeros fondos de pensiones, en 1743, el Scottish Minister's Widows' Pension Fund.

Es a finales del siglo XVII cuando arranca la historia moderna del seguro con la creación en Londres de la primera asociación profesional de aseguradores, cuyas primeras actuaciones se concentraron sobre el seguro marítimo y ya entrado el siglo XVIII se generalizaron al seguro de vida.

Durante el siglo XVIII, podemos decir que las Tablas de Mortalidad, vienen asociadas a los nombres de De Moivre, Bernoulli, Déparcieux, Euler, Dodson, Price y Morgan.

Deben destacarse las siguientes:

- Tabla Dodson para la Equitable Life (fundada en 1762), sobre una experiencia de mortalidad del periodo 1728-50.
- Tablas Holandesas de Struyck y Kersseboom, publicadas en 1740, y que tienen la virtud de, por primera vez, elaborar una tabla diferenciada de hombres y mujeres.
- Tabla Northhampton, elaborada por Price en base a la experiencia de mortalidad del periodo 1735-80 en dicha localidad inglesa. El libro de Price, consultor de Equitable Life, fue considerado manual de referencia hasta 1820.
- Tabla sueca Wargentin, sobre una experiencia de mortalidad del periodo 1755-63.

Para el siglo XIX y principios del XX pasamos a hacer un resumen siguiendo al profesor Lasheras (1948), sobre las experiencias extranjeras y reservamos el siguiente epígrafe como especifico para la experiencia española.

a.- Experiencia inglesa

El actuario Mr. Finlaison, en 1829, elaboró unas tablas de mortalidad por encargo del gobierno inglés. En 1834, el actuario A. Morgan construyó la primera tabla inglesa con datos de compañías de seguros correspondientes al período 1762-1828. Esta experiencia, fue mejorada y perfeccionada en 1843 por un comité de actuarios, con la publicación de las Tablas de las Diecisiete Compañías Inglesas (Combined Experience) con datos del período 1762-1837 (unos 84000 datos aproximadamente correspondientes a las diecisiete compañías inglesas más importantes). De esta época son dos trabajos muy relevantes debidos a Gompertz (1825) y a Makeham (1860). Respecto de este último hemos de decir que la mayoría de tablas de mortalidad hasta por lo menos mediados del siglo XX, se ajustaron a dicha formula.

En 1869 el Instituto de Actuarios de Londres y la Facultad de Actuarios de Escocia, publicaron cuatro tablas (las Healthy Males o $H^{m,}$ las Healthy Females o $H^{f,}$, las Diseased Males and Females o D.M.F., y las Special Risks o S.R) basadas en los datos de 10 compañías inglesas y otras 10 escocesas, que fueron de gran profusión. Más tarde, en 1903, el Instituto de Actuarios de Londres publicó las “The British Offices Life Tables, 1893” con datos de 60 compañías y un largo período de observación.

También, en cuanto a tablas generales de mortalidad, además de algunas experiencias en el siglo XVII, en 1859 se publicó la tabla de mortalidad general correspondiente al período 1849-54 con datos de 63 distritos importantes de Inglaterra y Gales, las cuales se perfeccionaron y mejoraron para los períodos 1881-90 y 1881-1900, tablas que se conocen con el nombre genérico de Heathy English Life Tables. Durante el soglo XX han ido elaborándose diferentes English Life Tables (ELT), incluso con criterios distintos, añadiendo a cada una uno o dos dígitos para diferenciarlas, así por ejemplo en el famoso libro de Benjamín and Pollard (1989) aparecen ejemplos de las tablas ELT13.

b.- Experiencia francesa

En 1860 se publicó en Francia la denominada Tabla del Comité de las Tres Compañías (L´Union, La National y Assurances Generales). En 1874, fue el actuario De Kertanguy quien publicó una tabla de mortalidad basándose en la experiencia de la Compagnie d´Assurances Generales y, en 1879, los actuarios Achard y Charlton construyeron otra basándose en datos de los pensionistas del estado francés. Más tarde, en 1887, el comité ya citado de las tres compañías construyó las famosas tablas de Assurés Francais (A.F.) y de Rantiers Francais (R.F) ajustadas por el método de Woolhouse, que poco después, en 1895, se ajustaron por la fórmula de Makeham. Finalmente, en 1899, se decidió la construcción de dichas tablas ajustadas a la fórmula de Makeham pero a un período más largo con comienzo en 1819, fecha de la fundación de la compañía más antigua de las tres, y final en 1898, las que fueron publicadas en 1900, con el nombre antiguo de Tablas del comité de las Tres compañías, (las AF para seguros en caso de muerte, las RF para seguros en caso de vida y rentas vitalicias) de gran profusión mundial y que prácticamente fueron usadas hasta los años 70 del s. XX. De finales del s. XIX es también muy importante el trabajo del actuario francés M. A. Quiquet (1893).

c.- La experiencia americana y australiana

La primera tabla con experiencia americana data de la segunda mitad del siglo XIX, y fue construida por Mr. H. Shephard actuario de la Mutual Life Insurance Company, basandose en los datos correspondientes a 16 años de dicha entidad. Mas tarde en 1881, Mr. Levi W.

Meech, dirigió la publicación de la Thirty American Offices Tables, basada en las experiencias de 30 compañías estadounidenses.

A principios del s. XX, el Departamento de Estadística de los Estados Unidos de América publicó una tabla de mortalidad general basada en el Censo de 1910 y las defunciones de los años 1909, 1910 y 1911, trabajo que dirigió J.W. Glover y que más tarde se completó diferenciando por razas, nacionalidad, núcleos urbanos etc. También, la Sociedad Actuarial Americana de Canadá y USA publicó en 1918 una tabla de mortalidad con las experiencias de compañías de seguros de América del Norte. Hoy en día son innumerables las publicaciones sobre tablas de mortalidad general y especificas existentes.

Las aportaciones australinas importantes aparecen a partir de 1950 con las famosas tablas ajustadas por el criterio de Helligman y Pollard (1980).

d.- Otras experiencias

También hay otras experiencias alemanas, del imperio austro-húngaro, belgas, de la república austro-húngara, e italianas.

Como resumen del siglo XX, podemos decir que:

a) Se produce, al menos en los países occidentales, una generalización en cuanto a publicaciones demográficas por los principales Institutos Oficiales de Estadística así como de Tablas de Mortalidad de la población general, así como por parte de algunos Institutos de Actuarios y Facultades. Por otra parte, las principales empresas (en general multinacionales) tanto de seguros como de reaseguros empiezan a elaborar sus propias tablas de colectivos de asegurados. Todo ello ha sido posible y potenciado, gracias a los equipos informáticos hoy en día disponibles.

b) Hasta el primer tercio del siglo XX, prácticamente, todas las tablas de mortalidad de uso en aplicaciones al seguro se ajustaron por la formula de Makeham (1860), y que a partir de la segunda mitad hay un cambio sustancial, debido a la aparición de otros criterios de ajuste y del extraordinario desarrollo de los sistemas informáticos que han revolucionado y siguen revolucionando las múltiples posibilidades de aplicar diferentes criterios de ajuste.

Hoy en día la famosa fórmula de Makeham (1860) basada en la ley de Gomptertz (1825), que se ajustaba por la formula de King y Hardy podemos decir que ha quedado obsoleta, con la incorporación de las funciones generalizadas de Gomperz-Makeham, las nuevas leyes de Helligman y Pollard (1980), los ajustes basados en funciones kernel (véase por ejemplo Ayuso, Guillen y otros (2001)), y las funciones splines (McCutcheon, J. J. (1981)), todo ello potenciado por el gran desarrollo de los equipos y softwares informáticos. Hay tantas aportaciones y en tantos sentidos que es imposible referenciar en un trabajo como éste.

Finalmente analizamos lo que podemos considerar como experiencia española en cuanto a tablas de mortalidad.

3.- Tablas De Mortalidad Españolas

A continuación analizamos la experiencia española a partir del s.XVII muy esquemáticamente en cuanto a los s. XVIII y XIX, y profundizando más en el s. XX.

Retrotrayéndonos a la ultima referencia histórica que hicimos sobre España, se puede constatar como, a partir del s. XVII, el declive del imperio español conlleva el desinterés por toda investigación estadística y hay que esperar hasta el s. XVIII, con el reinado de Carlos III, para encontrar nuevos estudios censales, que fueron continuados bajo los reinados de Carlos IV y Fernando VII (ya en el s. XIX). De esta época famosos fueron los censos del conde de Aranda (1768), de Floridablanca (1787), y de Godoy (1797) (Barrionuevo, 2001).

Dos hitos importantes hemos de destacar hasta finales del XIX: en 1852 se crea, por la Sociedad Económica Matritense, la primera cátedra de Estadística de España, que es ocupada por José María Ibáñez; y, en 1856 se funda la Comisión de Estadística del Reino, germen del actual Instituto Nacional de Estadística que, en 1877, junto con el hoy desaparecido Instituto Geográfico y Estadístico, elaboraron con los datos del Censo de 1877 y los fallecimientos ocurridos desde 1878 a 1882, una tabla de supervivencia de los habitantes de España, con grandes dificultades de consecución y tratamiento de datos, así como metodológicas, y ajustada por métodos gráficos.

En cuanto al s. XX, consideramos tres subapartados: por una parte las publicaciones oficiales del INE, por otra las publicaciones e intentos desde un punto de vista actuarial, y finalmente lo que denominamos punto de inflexión español.

A- Tablas de Mortalidad General: Publicaciones del INE

El Instituto Nacional de Estadística, tal y como hoy lo conocemos (cuyos predecesores como hemos dicho fueron la Comisión de Estadística del Reino, y el Instituto Geográfico y Estadístico) fue creado oficialmente por la Ley de 31 de diciembre de 1945 (BOE 03/01/1946) con la misión de elaborar y perfeccionar las estadísticas demográficas, económicas y sociales ya existentes, la creación de otras nuevas y la coordinación con los servicios estadísticos de las áreas provinciales y municipales.

La información en cuanto a tablas y publicaciones del INE es muy completa, la mayoría de las mismas pueden consultarse "on line" en la pagina web del mismo (www.ine.es), y entre las publicaciones en papel respecto a tablas de mortalidad, aunque todas a las que hemos aludido en el texto están publicadas, sólo incluimos en la bibliografía adjuntada la correspondiente a las Tablas de Mortalidad de la Población Española de 1996-97 (INE (1999)).

- Las primeras Tablas de Mortalidad de la población española del INE fueron publicadas en 1945 con datos del Censo de 1930 y los fallecimientos de los años 1930 y 1931, no tomándose los datos del Censo anterior de 31-12-1940, por su proximidad con la Guerra Civil (1936-39), pues eran previsibles muchos errores en los datos y una mortalidad mucho mayor que lo normal.
- En 1952, se publicaron las tablas de mortalidad correspondientes a 1940 y, además, aprovechando los Censos de 1900, 1910 y 1920, se elaboraron las correspondientes a esos períodos, con lo que se llenó el vacío que padecíamos con referencia a la primera cuarta parte del siglo.
- En 1960 aparecieron las tablas de 1950 y en 1977 las tablas (completas)de los años 1960 y 1970, si bien previamente, en 1963, se publicaron las tablas abreviadas de 1960.
- A partir de entonces el INE ha publicado tablas de mortalidad cada cinco años aprovechando los padrones municipales quinquenales, además de los censales (cada 10 años). A partir de 1975 se realizan cambios conceptuales como por ejemplo en el de

nacidos vivos, así como de algunos criterios en cuanto a las suavización de las series, adaptándolos y aproximándolos mas a la metodologías europea y americana.

- En 1981 aparecen las tablas de 1975-76, en 1988 las del 1980-81, en 1991 las del período 1985-86, y en 1993 las de 1990-91.

Con posterioridad ha habido un nuevo cambio en la periodificación de las publicaciones del INE sobre tablas de mortalidad, apareciendo en 1998 las Tablas de mortalidad del 94-95, utilizando las defunciones de dichos dos años pero no los datos de población del padrón correspondiente a 1995, sino los proyectados a partir del Censo de 1991, y en 1999 las del período 1996-97 siguiendo dicho criterio. Pero en general estas no se ajustaron ni a la ley de Makeham ni a otras posteriores de tipo actuarial, sino suavizadas en general por el método de las diferencias variantes (INE, 1999).

También, muy recientemente, a través de los Institutos de Estadística de algunas Comunidades Autónomas se han calculado tablas de mortalidad referidas a algunas de las Comunidades, y el propio INE en 1999 publico las Tablas de Mortalidad de la población española de 1994-95, con resultados por Comunidades Autónomas.

B- Publicaciones de otros Organismos Públicos, Semipúblicos y Privados, ajustadas con criterios actuariales

Así como la experiencia del INE, podemos considerarla adecuada comparada con la de otros países europeos, no podemos decir lo mismo de las publicaciones con criterio actuarial para utilizar por las compañías de seguros, lo que se acentuó muchísimo a partir de los años treinta del siglo pasado.

- En 1911 la entonces Comisaría General de Seguros (más tarde Dirección General de Seguros y hoy Dirección General de Seguros y Planes de Pensiones), intentó construir una tabla de mortalidad ajustada a la ley de Makeham, aunque no con experiencia española, y así se obtuvieron las tablas de mortalidad de los hermanos Puyol Lalaguna (1911).

  Conviene puntualizar que la primera ley de Seguros Privados en España data del 14 de mayo de 1908. Con anterioridad, las compañías de seguro que operaban en España disponían de una amplia libertad en el establecimiento de sus tarifas y reservas que dependían, en gran parte, de las indicaciones establecidas en los propios estatutos de constitución de las compañías. A finales del siglo XIX se había incorporado una normativa escasa que se limitaba a la obligación del pago de algunos impuestos, el depósito de garantía exigido por el Artículo 32 de la ley de Presupuestos de 5 de agosto de 1893 y del Artículo 43 de la Ley de Presupuestos de 30 de junio de 1895 o por el Real Decreto del Ministerio de la Gobernación de 27 de agosto de 1900, o el establecimiento de normativa cuando se creó el ramo de accidentes de trabajo en 1900. Esta libertad respecto a las reservas iba acompañada del bajo porcentaje de desembolso del capital social suscrito. Estos dos elementos, junto con el carácter especulativo con que se dirigió durante el siglo XIX la inversión de las escasas reservas en las compañías, explicarían la corta supervivencia de las compañías de seguro en esta etapa.

  La Ley de 1908 intentó paliar estas deficiencias estableciendo un estricto control sobre las compañías aseguradoras para garantizar los derechos de los asegurados. Esta normativa suponía que las compañías autorizadas estarían, a partir de estos momentos, sujetas a una serie de obligaciones respecto a la administración y a los propios asegurados.

El año 1908 es comúnmente denominado «el año de oro del seguro español» pues en él se promulgaron dos leyes de extraordinaria trascendencia: por una parte la ley de 27 de febrero de 1908 por la que se creó el Instituto Nacional de Previsión, y por otra la Ley del seguro privado de 17 de mayo 1908. En dichas leyes se ponen de manifiesto las operaciones de seguros, y también la profesión de actuario aunque entonces en España ni existía la carrera ni el Instituto de Actuarios.

La importancia de la primera ley la pone de manifiesto su artículo 15: «En la práctica de estas operaciones (a las que se le atribuirán las de renta vitalicia, diferida o personal, constituida a favor de personas de las clases trabajadoras) observará estrictamente el «Instituto Nacional de Previsión» las reglas técnicas del seguro. A este efecto, y debidamente asesorado por un Actuario de Seguros, con título profesional nacional o extranjero (atención a la automática homologación del título de Actuario sin necesidad a esperar a los trabajadores transfronterizos...), formulará el Consejo Patronato las tarifas de cuotas, con arreglo a la tabla de mortalidad que se considere preferible, de las utilizadas para el Seguro en caso de vida, mientras no tenga tabla nacional propia, y al tipo de interés que acuerde, no excediendo del 3,5%, con el recargo que se considere conveniente para constituir una reserva especial a los efectos de las fluctuaciones en la mortalidad y en el interés de las inversiones»

La Ley de 1908 fue desarrollada por el Reglamento, de 2 de febrero de 1912, (que estuvo vigente hasta la promulgación del nuevo Reglamento de Ordenación del Seguro Privado de 1 de agosto de 1985, hoy en día también modificado). En el artículo 99 del mismo se recomendaba la elaboración de unas tablas de supervivencia basadas en experiencias españolas y además, proporcionaba una lista de tablas de otros países (entre ellas las francesas AF y RF de 1895), que podían ser usadas por las compañías de seguros españolas o que operasen en territorio nacional, a la que haremos referencia mas adelante.

Podemos afirmar que, como consecuencia de ambas leyes, aparecieron las tablas de Puyol ya mencionadas y poco más tarde la regulación de la carrera de actuario y la de la profesión de actuario y la creación del instituto de actuarios (Escuder J y R, y Vegas (2006)).

- Más tarde, en 1927, el (entonces) Consejo Superior de Trabajo Comercio e Industria, publicó unas tablas debidas a Mariano Fuentes Martiáñez (1927), donde aparecía la sobrevivencia o supervivencia, vida media y vida probable, basadas en los Censos de 1910 y 1920 y en las defunciones del período 1908-1923 (esto es defunciones del período intercensal, mas tres años anteriores y tres posteriores). En dicha publicación no se consideraron los datos de 1918 por la terrible epidemia de gripe habida que cercenó muchas vidas en nuestro país. En dicha tabla, aunque se intentó ajustar a la ley de Makeham se introdujeron muchos trucos y artificios que hicieron dudar de su fiabilidad.
- El episodio cruento de nuestra Guerra Civil (1936-1939), modificó sustancialmente la mortalidad del periodo, y produjo cambios institucionales y legislativos, que perduraron casi durante cuarenta años de aislacionismo, aunque en los últimos diez se produjeron aperturas.
- El profesor Lasheras (1948) relata un nuevo intento de construir unas tablas actuariales en 1945 de utilidad para la Seguridad Social por parte de la Asesoría Técnica de la Dirección General de Previsión (Ministerio de Trabajo).
- Hasta 1954 no hubo modificación sustancial de la ley de 1908. La nueva Ley de Seguros de 16 de diciembre de 1954, no proporcionó los efectos esperados. Sus cambios y

adaptaciones no fueron muy operativos ya que si bien, en su artículo 4º prescribía la necesidad de presentación de "las bases técnicas, tarifas, pólizas y contratos que se propongan utilizar en sus operaciones", sin embargo, como esta ley no tuvo reglamento propio (que siguió siendo el de 1912), no quedó regulado lo que se entenderían por "bases técnicas" en la práctica aseguradora. Como consecuencia de la misma, si que hubo un nuevo intento de construir unas tablas actuariales con experiencia española como detalla el profesor Lóbez (1966), pero al igual que el de 1945, sin llegar a resultados verdaderamente efectivos.

Relacionado con lo anterior, es de especial importancia citar la Orden Ministerial de 8 de febrero de 1961, que devino fácticamente en reglamento de la ley de 1954 (hasta que por fin en 1985 se publicó el Reglamento de la nueva Ley sobre Ordenación del Seguro Privado 33/1984, de 2 de agosto de 1984, mediante Real Decreto 1348/1985, de 1 de agosto, que derogó dicha O. M,, así como el ya caduco Reglamento de 1912).

Esta O. M.: liberalizó (de "aquella forma") el mercado asegurador siguiendo los vientos "liberalizadores" que inspiraban la política económica de la época. En lo que se refiere a las tablas de mortalidad, la exposición de motivos de la orden establecía que "la ausencia transitoria de tablas de mortalidad, establecidas sobre la experiencia española, obliga a aceptar todas aquellas que reúnan las condiciones de idoneidad técnica, siempre que las empresas que soliciten su empleo demuestren a satisfacción de la Dirección General encargada de la vigilancia y tutela del seguro privado, que su adopción es posible por resultar de los datos estadísticos oportunamente aportados que la mortalidad real en los colectivos españoles es igual o menor que la de las tablas propuestas en los seguros en caso de muerte y, caso necesario, en los seguros para el caso de vida".

De esta forma, el apartado tercero de la mencionada Orden Ministerial establecía: *"Las entidades operantes en el ramo de vida podrán usar, transitoriamente, para sus cálculos, cualquier tabla de mortalidad, siempre que demuestren inequívocamente ante el Ministerio de Hacienda, mediante estudios estadísticos, que la mortalidad real española es igual o menor que la contenida en las tablas de mortalidad que se pretendan utilizar en los seguros en caso de muerte o igual o superior en caso de vida"*.
La realidad fue que la mayoría de las compañías siguieron aplicando el articulo 99 del Reglamento de 1912, en el que se decía explícitamente que: *"Las Tablas de Supervivencia y de Mortalidad, para el calculo de tarifas en caso de seguros de vida y de reservas matemáticas, autorizadas eran para los casos de muerte:*

- *las AF (francesas),*
- *las HM inglesas (1902),*
- *las austrohúngaras de 1907,*
- *las de la experiencia americana,*
- *la alemana del colegio de Berlin MH,*

y para los de vida:

- *las RF (francesas),*
- *la de 17 compañías inglesas,*
- *la de la experiencia del gobierno británico,*
- *la de Carlisle,*
- *la de la experiencia americana.*

También en el mismo articulo se decía que *podrían emplearse otras basadas en experiencias particulares de las entidades aseguradoras o las deducidas de otras estadísticas de mortalidad, siempre que las primas netas que resulten de su aplicación queden comprendidas para todas las edades entre la más alta y la más baja de las que correspondan a las mismas edades con arreglo a las tablas anteriores*".

Tampoco en el ámbito académico se publicaron tablas de mortalidad propiamente españolas. Así, para los para estudiantes de Ciencias Económicas que en el quinto y último año cursaban la carrera actuarial (la carrera actuarial, hasta 1953 se cursaba en las Escuelas de altos Estudios Mercantiles (Escuder J. y R, y Vegas A (2006)) se utilizaban los siguientes dos libros de tablas: "Tablas Financieras y Actuariales" de Gil Peláez L. (1959) y el "Prontuario de Tablas Financieras y Actuariales"de Sánchez Valverde (1956). Tablas que en la parte actuarial eran fundamentalmente las propias francesas, inglesas e italianas aludidas en el artículo 99 del Reglamento de 1912.

Hoy en día las principales empresas de seguros multinacionales y/o reaseguradoras, y también algunas españolas sobre toda en el ámbito de banca-seguros, elaboran sus propias tablas, a partir de sus propios registros, a las que se denominan tablas de colectivos asegurados en contraposición de las generales o construidas a partir de la población general, pero de este aspecto junto a otros muy positivos ocurridos en España últimamente, los analizamos a continuación, como nuestro punto de inflexión hacia nuestra identificación europea.

C- Cambio o inflexión en España a partir del periodo 1970-1980

Como acabamos de constatar, en España aparte de las publicaciones del INE en cuanto a mortalidad general y no ajustadas por criterios actuariales, verdaderamente no hemos dispuesto de tablas especificas de mortalidad obtenidas de datos de la población española asegurada de uso por las compañías aseguradoras, sino que más bien hemos utilizado tablas extranjeras.

Puede que todo ello sea debido al tamaño de nuestras entidades aseguradoras y a su fuerte dependencia de las de las grandes multinacionales aseguradoras y reaseguradotas, y a nuestras propias circunstancias sociales, políticas económicas y jurídicas (recuérdese además, que el reglamento de 1912 estuvo jurídicamente vigente hasta 1985).

Sin embargo a partir del período 1970-1980 hay un cambio sustancial facilitado por el nuevo marco jurídico a partir de la democracia española, que pasamos a exponer. Entre dichos hitos destacamos:

- Los cambios legislativos debidos a nuestra democracia y por la entrada un poco más tarde (1986) en la hoy Unión Europea por las adaptaciones que tuvimos y que tenemos que hacer a sus normativas.

Así, citamos entre otras: la Ley 8 de octubre de 1980 que desarrollaba el contrato del seguro; la nueva Ley 33/1984, de 2 de agosto sobre Ordenación del Seguro Privado, y su decreto desarrollador publicado mediante Real Decreto 1348/1985 de 1 de agosto, que por fin derogó el de 1912; la nueva ley denominada de Ordenación y Supervisión de los Seguros privados de 8 de Noviembre de 1995, y su reglamento desarrollador de 29 de noviembre de 1998 (que también modifica parcialmente el de 1985). Con posterioridad se han publicado como nuevas Ordenes Ministeriales, Decretos, y Textos Refundidos, ya que la legislación actual española se moderniza y modifica casi continuamente por motivos de actualización y por adaptación a las normativas emitidas por la Unión Europea. La última modificación del Reglamento de Seguros vigente, es el reciente Real Decreto 239/2007 de 16 de febrero**.**

En cuanto a las tablas de mortalidad y su utilización hemos de destacar que el Artículo 34 del Reglamento de Ordenación y Supervisión de los Seguros Privados (Real Decreto 2486/1998, de 20 de noviembre) establecía en el apartado 1 las características que deben reunir las Tablas de mortalidad, de supervivencia y de invalidez para ser aplicadas en el Seguro de Vida en España. Así, se establece con carácter general que deben *"estar basadas en experiencia nacional o extranjera, ajustada a tratamientos estadístico-actuariales generalmente aceptados" (apartado a), que "la mortalidad, supervivencia e invalidez reflejadas en las mismas deberán encontrarse dentro de los intervalos de confianza generalmente admitidos para la experiencia española" (apartado b), que "el final del período de observación considerado para la elaboración de las tablas no podrá ser anterior en más de veinte años a la fecha de cálculo de la provisión" (apartado c), que "cuando se utilicen tablas basadas en la experiencia propia del colectivo asegurado, la información estadística en la que se basen deberá cumplir los requisitos de homogeneidad y representatividad del riesgo, incluyendo sobre el mismo información suficiente que permita una inferencia estadística e indicando el tamaño de la muestra, el método de obtención de la misma y el período a que se refiere, el cual deberá adecuarse a lo previsto en el párrafo c) anterior" (apartado d) y que "en los seguros de supervivencia, deberán incorporar el efecto del tanto de disminución de la mortalidad considerando una evolución desfavorable de la misma, salvo que el mismo haya sido tenido en cuenta en el cómputo del período de observación a que se refiere el párrafo c) anterior" (apartado e).*

Sin embargo, la Disposición Transitoria Segunda de dicho Reglamento de Ordenación y Supervisión de los Seguros Privados establecía también en su apartado quinto, que *"no obstante lo dispuesto en el párrafo c) del apartado 1 del articulo 34 de este Reglamento, hasta tanto así se declare por la Dirección General de Seguros por haberse contrastado la validez de nuevas tablas de final de período de observación más reciente, a partir de la entrada en vigor de este Reglamento podrán utilizarse para seguros de supervivencia las tablas GRM80 y GRF80 con dos años menos de edad actuarial y en seguros de fallecimiento las tablas GKM80 y GKF80.*

La vigencia de esta disposición transitoria fue corta, como constatamos a continuación por la publicación de las primeras tablas basadas en experiencia española que fueron las tablas actuales PERM Y EPRF 2000, a las que nos referiremos inmediatamente.

- Los trabajos de investigación realizados por los profesores universitarios Eliseo Navarro (1982) y Eugenio Prieto y Javier Fernández Plasencia (1994), que además de su importancia estimularon la incipiente investigación en materia actuarial de muchas universidades españolas y empresas, con tesis doctorales y trabajos de investigación importantes.

El trabajo de Eliseo Navarro, editado y premiado por la institución Mapfre, titulado: "Tablas de Mortalidad de la Población Española 1982: Metodología y Fuentes" (1991), en el que siguiendo a los profesores Forfar, D. O, MacCutcheon, J. J. y Wilkie A. D.,(1988), abandona el criterio único de ajuste a la formula de Makeham., introduciendo los ajustes a diferentes funciones denominadas funciones generalizadas de Gompertz-Makeham.

El trabajo de Eugenio Prieto y Javier Fernández Plasencia, publicado en 1994 por el Servicio Actuarial de Unespa y titulado "Tablas de Mortalidad de la Población Española de 1950 a 1990. Tabla Proyectada del año 2000", donde se aborda el problema del ajuste, suavización o graduación de las misma a la fórmula de Makeham (al menos para el período central de las mismas) a partir de las tablas de INE. En este aspecto debemos recordar que las tablas de mortalidad publicadas por el INE, no estaban ajustadas a ninguna ley de mortalidad de tipo actuarial, sino suavizadas por el método de las diferencias variantes.

Y finalmente,

- El Protocolo de colaboración de la Dirección General de Ordenación de la Seguridad Social, Dirección General de Seguros (DGS y PP), UNESPA, ICEA, el INE y el Instituto de Actuarios españoles, para la construcción de las Tablas de mortalidad de asegurados españoles, que cristalizó finalmente en la construcción y publicación por dicha Dirección General de las Tablas Generacionales Españolas del año 2000 (Resolución de 3 de octubre de 2000).

Así, en la exposición de motivos de dicha Resolución de 3 de octubre de 2000 de la Dirección General de Seguros y de Planes de Pensiones, dando cumplimiento a lo previsto en el número 5 de la disposición transitoria segunda del Reglamento de Ordenación y Supervisión de los Seguros Privados, aprobado por Real Decreto 2486/1998, de 20 de noviembre, en relación con las tablas de mortalidad y supervivencia a utilizar por las entidades aseguradoras, se prescribe:

“Primero. Declarar, en aplicación de lo previsto en el número 5 de la disposición transitoria segunda del Reglamento de Ordenación y supervisión de los seguros privados, aprobado por Real Decreto 2486/1998, de 20 de noviembre, la no admisibilidad de la utilización de las tablas GRM80 y GRF80 corregidas con dos años menos de edad actuarial para garantías de supervivencia.

Segundo. Hacer públicas, mediante su inclusión en anexo a la presente resolución, las tablas denominadas PERM/F-2000 (Tablas Generacionales Españolas de Supervivencia Masculina/ Femenina), que podrán utilizarse por cumplir los requisitos exigidos en el artículo 34 del Reglamento de Ordenación y supervisión de los seguros privados, en los siguientes términos:

- *Las Tablas PERM/F-2000P serán de aplicación a la nueva producción que se efectúe desde la entrada en vigor de la presente resolución, así como a las incorporaciones que se produzcan a partir de la misma fecha a pólizas colectivas ya en vigor.*

- *Las Tablas PERM/F-2000C serán de aplicación a la cartera de pólizas en vigor a la misma fecha, debiendo efectuarse la primera dotación correspondiente a la adaptación a estas tablas en el presente año 2000, y encontrarse la misma concluida en el plazo de 13 años a contar desde el día 1 de enero de 2001, todo ello de conformidad y con arreglo a los criterios que se deducen de lo previsto en el número 4 de la disposición transitoria segunda del Reglamento”.*

Esta resolución entró en vigor el día 15 de octubre de 2000, a partir del cual se hace uso exclusivamente de las mencionadas tablas de experiencia española para los seguros de vida en España. Es la primera vez que en España no tenemos que recurrir obligatoriamente a tablas de otros países para calcular nuestras primas, aunque algunas compañías también utilizan las GR95 puesto que no han transcurrido veinte años desde su publicación, ya que el apartado c del Artículo 34 del Reglamento de Seguros (Real Decreto 2486/1998, de 20 de noviembre) establece que “el final del período de observación considerado para la elaboración de las tablas no podrá ser anterior en más de veinte años a la fecha de cálculo de la provisión”.

4.- A modo de Conclusión

Evidentemente, en España ha habido cierto retraso en cuanto a la confección de tablas de mortalidad con criterios actuariales pero hoy, teniendo en cuenta lo que acabamos de indicar en el subepígrafe anterior, podemos decir que debido:

- por una parte a la publicación de las Tablas Generacionales Españolas PERM PERF 2000, por la Dirección General de Seguros y de Planes de Pensiones, organismo público que goza de personal muy especializado en materia de seguros y pensiones. Téngase en cuenta que es la primera vez que en España no tenemos que recurrir a tablas de otros países para calcular nuestras primas (aunque las GR95 todavía están también vigentes)
- por otra a los nuevos equipos de investigación universitarios, y sus publicaciones, y
- finalmente, por otra, a las exigencias y requisitos de obligado cumplimiento de la recientes directrices en materia de seguros y sobre todo de solvencia marcadas por la Unión Europea, que obligan a estudios muy profundos (y así tenemos constancia que algunas instituciones de banca seguros españolas han construido tablas actuariales propias) y sofisticados y a contratar personas de gran cualificación, fundamentalmente titulados procedentes de la titulación actuarial, desempeñando tareas directivas y de investigación,

Podemos afirmar por lo tanto, que nuestro retraso crónico si no ha desaparecido completamente, se ha mitigado muchísimo respecto de los países e instituciones destacadas en el contexto mundial.

4.- Apéndice Matemático

En este apéndice matemático incluimos muy resumidamente una relación de las principales leyes y formulas de ajuste utilizadas en las Tablas de Mortalidad.

- <u>*De Moivre*</u> (1667-1754) alrededor de 1693 formuló la siguiente ley de mortalidad: $l_x = 86 - x$

basado en los cálculos de Halley sobre la mortalidad de la ciudad de Londres en la que partía de una tabla de 1.000 efectivos y de razón decreciente en % e igual a 10. Por ello De Moivre partiendo de unos 100 (exactamente 86 efectivos, w=86) y considerando una razón fija q = 1% formuló la citada ley. El método de Halley fue uno de los precursores remotos del actual método de censo según Insolera (1950).

- <u>*D'Alembert*</u> en 1765 , basándose en los censos de la ciudad de Londres propuso la siguiente:

$$l_x = a + b(x-45) + c(x-45)^2 + d(x-45)^3 + e(x-45)^4 + f(x-45)^5$$

- <u>*Babbage en 1823*</u> formuló:

$$l_x = a + b.x + c.\frac{x.(x-1)}{2}$$ (siendo a positivo, y b y c negativos).

- <u>*Gompertz,*</u> como ya hemos estudiado, en 1825, propuso:

$$\mu_x = b.c^x \qquad \rightarrow \qquad l_x = k.g^{c^x}$$

- <u>*Makeham*</u> en 1867 reformuló la ley de Gompertz, mediante:

$$\mu_x = a + b.c^x \qquad \rightarrow \qquad l_x = k.s^x.g^{c^x},$$

la que se conoce como primera ley de Makeham. También existe la llamada segunda ley de Makeham que es:

$$\mu_x = a + dx + b.c^x \qquad \rightarrow \qquad l_x = k.s_1^x.s_2^{x^2}.g^{c^x}.$$

- *Lazarus* en 1867 formuló, su ley

$$\mu_x = a + b_1.c_1^x + b_2.c_2^x \qquad \rightarrow \qquad l_x = k.s^x.g_1^{c_1^x}.g_2^{c_2^x}.$$

- que fue generalizada por *Jamse* en el siguiente sentido (orden k>2):

$$\mu_x = a + b_1.c_1^x + b_2.c_2^x + b_3.c_3^x + ... + b_k.c_k^x \quad . \quad l_x = k.s^x.g_1^{c_1^x}.g_2^{c_2^x}...g_k^{c_k^x}.$$

- *E. Sang* matemático y actuario escocés nacido en 1805 y fallecido en 1890, fue miembro fundador de la Facultad de Actuarios de Escocia y propuso la primera ley exponencial para la función biométrica lx (no para μ_x) , esto es:

$$l_x = a + b.c^x.$$

- *Thiele* en 1872 propuso la primera ley basada en tres funciones una para cada rango de edad que denominó: niñez, intermedio y vejez, mediante la siguiente función:

$$\mu_x = a_1 e^{b_1 x} + a_2 e^{-(1/2)b_2(x-c)^2} + a_3 e^{b_3 x}.$$

- *Emile Dormoy,* actuario francés (1829 -1891), entre 1874 y 1878 y propuso las siguientes leyes de mortalidad:

1ª Ley de Dormoy:

$$\mu_x = a \text{ (constante) } \quad l_x = k.S^x \quad {}_tP_x = S^t \quad {}_{/t}q_x = 1 - S^t.$$

2ª Ley de Dormoy

$$\mu_x = a.x \qquad l_x = k.S_1^x.S_2^{x^2} \qquad {}_tP_x = S_1^t.S^{t^2+2tx} \qquad {}_{/t}q_x = 1 - S_1^t.S^{t^2 2tx}.$$

3ª Ley de Dormoy

$$l_x = k.S_1^x.S_2^{x^2}.S_3^{x^3}.$$

- *Risser*, ya en el siglo XX (concretamente en 1908) propuso la siguiente ley:

$$l_x = S_1^x.\left(S_2^x\right)^{c^x}.g^{c^x}.$$

- *Perks* que en 1932 basándose en un trabajo previo de Karl Pearson, propuso la siguiente ley:

$$\mu_x = \frac{a + bc^x}{kc^{-x} + 1 + dc^x}.$$

- Con posteriridad, los profesores Forfar, D. O, MacCutcheon, J. J. y Wilkie A. D., que en 1988 publicaron en el Journal of the Institute of Actuaries, 115, 1-149, el trabajo titulado “On Graduation by Mathematical Formula”, que aparte de ser una obra de referencia obligada para cualquier investigador, en ella se revisan las bases estadístico-matemáticas de la graduación, y se introducen las funciones generalizadas de Gompertz-Makeham como una nueva formación para resolver problemas concretos de graduación. La estructura general de la misma es la familia de funciones generalizadas de tipo exponencial y polinómico para graduar el tanto instantáneo de mortalidad:

$GM^{r,s}(x) = GM(r,s) = P^1_{r-1}(x) + \exp\{P^2_{s-.1}(x)\}$

donde $P^1_{r-1}(x)$ y $P^2_{s-.1}(x)$ son polinomios en x de grado r-1 y s-1 respectivamente, esto es:

$$\mu_X = GM(r,s) = a_0 + a_1.x + a_2.x^2 + ... + a_{r-1}.x^{r-1} + \exp\{b_0 + b_1.x + ... + b_{s-1}.x^{s-1}\},$$

donde en general $0 \le r, s \le 5$.

Las constantes o parámetros del modelo generalizado deben obtenerse por criterios estadísticos procurando que el modelo elegido sea lo más parsimonioso posible, esto es contenga el menor número de parámetros. El poder graduar con diferentes modelos permite compararlos y seleccionar, entre aquellos de propiedades estadísticas parecidas en cuanto a la bondad de los ajustes pertinentes, los más parsimoniosos como hemos indicado.

Dichos autores también consideraron la formulación Logit Gompertz-Makeham Generalizada, dada por la expresión Logit de la formula generalizada de Gompertz-Makeham que es:

$$\mathrm{LGM(r,s)} = \frac{GM^{r,s}(x)}{1+GM^{r,s}(x)}$$

- Los actuarios australianos Heligman L, and Pollard J.H , en su trabajo "The Age Pattern of Mortality" publicado en 1980 en el Journal of the Institute of Actuaries nº 107 pp 49-80, introducen tres tipos de nuevas funciones para representar la mortalidad, que se ajustaron muy bien a la experiencia australiana, estas son:

a.- tipo $$q_x = A^{(x+b)^C} + E\exp\{E(\ln x - \ln F)^2 + \frac{G \bullet H^x}{1+G \bullet H^x}$$

b.- tipo II $$q_x = A^{(x+b)^C} + E\exp\{E(\ln x - \ln F)^2 + \frac{G \bullet H^x}{1+K * G \bullet H^x}$$

c.- tipo III $$q_x = A^{(x+b)^C} + E\exp\{E(\ln x - \ln F)^2 + \frac{g \bullet H^x}{1+G \bullet H^{x^K}}$$

y en general válidas para todo el intervalo de edades (0,w), y que tienen la propiedad que sus parámetros pueden interpretarse biométricamente.

- Finalmente indicamos la posibilidad de ajuste o graduaciones mediante funciones núcleo (Ayuso, Corrales, Guillen, Marin y Rojo (2001)), y splines. McCutcheon J.J.(1981).

Bibliografia

AYUSO, CORRALES, GUILLEN, MARIN Y ROJO (2001). *Estadística Actuarial Vida*. Edicions Universitat de Barcelana (EUB).

BARRIONUEVO DOLMOS AUGUSTO (2001). *Las estadísticas históricas en el INE*. Revista Fuentes Estadísticas. nº 50.

BENÍTEZ DE LUGO, LUIS, (1955). *Tratado de Seguros.* Volumen 1, pag 114.

Instituto Editorial Reus.

BENJAMIN, P. AND POLLARD. J. (1989). *The Analysis of Mortality and Other Actuarial Statistics*. Ed. Heinemann. London.

ESCUDER J, ESCUDER R Y VEGAS (2006)). *La Titulación de Licenciado en Ciencias Actuariales y Financieras en España en el nuevo contexto del Espacio Europeo de Enseñanza Superior*. 5º Congreso Argentino de Actuarios y 8º Congreso Panamericano de Actuarios. Buenos Aires, 18 al 20 de octubre de 2006.

FORFAR, D. O, MACCUTCHEON, J. J. Y WILKIE A. D.,(1988). *On Graduation by Mathematical Formula* JIA 115, 1-149.

FUENTES MARTIAÑEZ M.(1927). *Tablas de mortalidad, supervivencia, vida media y vida probable.* CSTCI Madrid.

GOMPERTZ, B.(1825). *On de nature of the function Expressive of law of Human Mortality and on a New Mode of determining the Value of Life Contingencies*. Philosophical Transactions of the Royal Society. vol. 115, pp 513-585.

HELIGMAN L. AND POLLARD J. H (1980). The Age Pattern of Mortality"

JIA nº 107, pp 49-80.

INE (1999). *Tablas de Mortalidad de la Población Española, 1996-97.* Instituto Nacional de Estadística. Madrid.

INSOLERA, F. (1950). *Curso de Matemática Financiera y Actuarial*. Ed. Aguilar (traducción de la edicion italiana realizada en 1937, bajo los auspicios de la sociedad Reale Mutuadi Assicurazioni di Torino).

LASHERAS A. (1948). *Matemática Actuarial*. Ed. Dossat

LÓBEZ URQUÍA A. (1966). *Matemática financiera con Nociones de Calculo Actuarial*.

Barcelona (propiedad del autor).

MCCUTCHEON, J. J. (1987).*Experiments in graduating the data for the English Life Tables (nº 14).* Transactions of the Faculty of Actuaries. Vol. 40 (1)

MCCUTCHEON J. J. (1981).*Some Remarks on Splines*. Transactions of the Faculty of Actuaries, vol.37 pp.421-438

MAKEHAM, W. M. (1860).*On the Law of Mortality and the construction of Annuity Tables.*Journal of Institute of Actuaries vol 8, pp 301-310.

NAVARRO, E. (1991).*Tablas de mortalidad de la población española 1982: metodología y fuentes (1991).* Mapfre

PRIETO, E., FERNÁNDEZ PLASENCIA J. (1994).Tablas de Mortalidad de la Población Española de 1950 a 1990, y Tabla Proyectada del año 2000. UNESPA. Madrid.

PUYOL LALAGUNA , M. (1911).*Tabla de mortalidad española ajustada analíticamente.* Ed. Ricardo F. de Rojas, Madrid.

QUIQUET M. A.(1893). *Representation algebrique des tables de survies.* Bulletin de l´Institut des Actuaires Français. num 14.

SÁEZ DE JÁUREGUI SANZ L. M.(2007). *La igualdad efectiva de mujeres y hambres: aspectos relacionados con las tablas actuariales de supervivencia*. IAE: Actuarios: 2007, pp. 1 a 16 (Dossier).

SÁNCHEZ VALVERDE A. Y OTROS (1955 y 1956). *Prontuario de fórmulas y tablas financieras y actuariales (vol I y II).* Zaragoza (Vol. I, 1955 y Vol. II, 1956). Propiedad de los autores.

TEUGELS-SUNDT EDITORS (2004). *Encyclopaedia of Actuarial Science*. Ed Wiley. pp. 1139-1145.

VEGAS A. (1981). *Estadística. Aplicaciones Econométricas y Actuariales*. Ed. Pirámide.

Capítulo 20

Una historia sobre las relaciones entre eficiencia y aleatoriedad

JOSÉ JAVIER BUSTO GUERRERO
JESÚS MUÑOZ SAN MIGUEL
Universidad de Sevilla

Introducción

La teoría de la eficiencia informativa de los mercados financieros, tal como se formula en nuestro tiempo, surgió a mediados del siglo XX a partir del estudio por parte de los economistas sobre la gran cantidad de datos estadísticos que se habían acumulado durante treinta años y que no habían originado una interpretación económica adecuada. La descripción puramente probabilista del análisis de las variaciones bursátiles, por una parte, y la reflexión económica sobre el papel de las informaciones proporcionadas por los precios y la naturaleza de los sistemas de precios, por otra, se reunieron en los años 1960 para dar lugar a la teoría de eficiencia informativa de los mercados. Hayek (1945) dio una primera definición de la naturaleza de los mercados utilizando el concepto información transmitida por el sistema de precios, mientras que, a partir de la tesis de Bachelier (1900), los estadísticos buscaban caracterizar el comportamiento de los mercados bursátiles mediante técnicas aleatorias.

La extensión de su campo de aplicación ha sido inmensa. Los financieros profesionales instrumentalizaron el concepto de eficiencia en la gestión de sus activos, con la creación de gestiones mecanizadas o casi mecanizadas llamadas indiciales. El desarrollo del gran ahorro en EE.UU. a través de los fondos de pensiones aceleró la idea de eficiencia en los bancos y las sociedades de gestión, a la vez que las bolsas organizadas comenzaban a elaborar un número cada vez más grande de índices de representación de los mercados de acuerdo con los principios y las consecuencias de la teoría de la eficiencia. Más adelante, el concepto de eficiencia, se difundió aún más con la creación de los mercados de opciones sobre instrumentos financieros cada vez más complejos, a medida que los desarrollos matemáticos acompañaban y favorecían el crecimiento de estos mercados de capitales.

Las convulsiones intelectuales que originó la teoría de la eficiencia informativa fueron considerables. Era necesario construir una nueva manera de pensar los mercados, en contra de la intuición y no habitual, en relación a lo que dictaban los manuales clásicos de gestión financiera. Los cambios introducidos por el concepto produjeron modificaciones significativas de las prácticas profesionales. En este sentido, Bernstein (1992) no duda en afirmar que este nuevo enfoque de los mercados "parecía implicar modificaciones radicales en la manera en la que debería ser aplicado a sus responsabilidades profesionales". Fama (1991) piensa que el concepto de eficiencia "ha cambiado la visión que los profesionales llevan al mercado". En definitiva, existe un gran acuerdo en que la teoría de la eficiencia informativa ha supuesto una auténtica revolución epistemológica, que deriva de la introducción de los procesos aleatorios en la finanzas, y del abandono de los enfoques clásicos a favor de una perspectiva probabilista en el enfoque de las variaciones bursátiles.

El marco probabilista básico del concepto de eficiencia fue expuesto por primera vez por Bachelier (1900), después precisado por Osborne (1959), para ser posteriormente desarrollado por Samuelson (1965) y Mandelbrot (1966).

Con carácter general, las evoluciones de los activos financieros se modelan en la teoría financiera mediante exponenciales de procesos aleatorios que formalizan sus ecuaciones de comportamiento. El modelo tipo de las variables bursátiles es la ecuación

$$S(t) = S(0)\exp(\mu t + \sigma W(t)),$$

con $\mu \in \mathbb{R}$, $\sigma > 0$ y $W(t)$ un movimiento browniano estándar. Esta ecuación representa de manera formal la evolución de las trayectorias de los diferentes activos financieros. Si llamamos tasa de rentabilidad a

$$r(t) = \mu t + \sigma W(t),$$

esta representación hace la hipótesis fuerte de que la rentabilidad de un activo financiero cualquiera puede ser modelado por un proceso aleatorio de incrementos independientes y estacionarios, cuya ley marginal es una ley de Laplace-Gauss. Lo que significa, entre otras cosas, que sus características probabilistas están determinadas de forma única por la esperanza (la media) y la desviación típica. Por ello, toda estrategia de inversión se funda, en última instancia, sobre consideraciones que se basan en la pareja (μ,σ), rentabilidad y riesgo o esperanza matemática y desviación típica de la tasa de rentabilidad instantánea del activo considerado. Con esta concepción de las rentabilidades, aplicando estos útiles conceptuales para la elección de portafolios con futuro incierto, Markowitz, Tobin y Sharpe han fundamentado las bases de la teoría del portafolio.

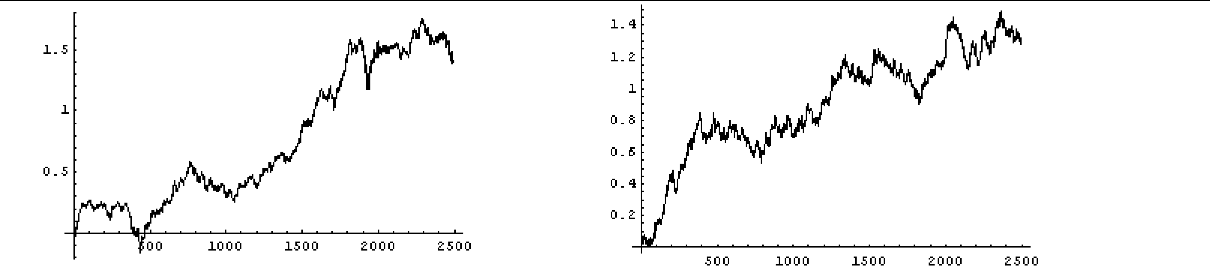


Figura 1: Tasa de rentabilidad del Ibex35 durante la década de los 90 frente a una simulación de un movimiento browniano con la misma media y varianza.

El primer aspecto a subrayar en esta modelación es la hipótesis de que el proceso aleatorio de la tasa de rentabilidad tiene incrementos independientes e idénticamente distribuidos (en adelante escribiremos i.i.d.). Esta hipótesis es muy cómoda para el cambio de escala de las rentabilidades bursátiles periódicas, puesto que en este caso $P(\Delta r_{t,\tau})$ es la distribución de probabilidad de las rentabilidades periódicas $\Delta r_{t,\tau} = \ln[S(t)/S(t-\tau)]$, entonces

$$P(\Delta r_{t,n\tau}) = P(\Delta r_{t,\tau})^{\otimes n} \quad , (*)$$

donde $\otimes$ representa el operador convolución. Desde el punto de vista probabilista el interés de la hipótesis i.i.d. es de cálculo. Desde el punto de vista económico, la independencia de las rentabilidades está relegada, como se verá, a la hipótesis de eficiencia informativa. La identidad de las distribuciones significa que las características económicas del fenómeno observado no cambian *demasiado* en el curso del tiempo.

En el caso de la modelación financiera estándar, la relación (*) se simplifica algo más, ya que el movimiento browniano presenta propiedades de invarianza por cambio de escala:

$$P(\Delta r_{t,\tau})^{\otimes n} = n^{1/2} P(\Delta r_{t,\tau})$$

Se puede interpretar esta relación como una *propiedad fractal de autosemejanza*: el proceso de las tasas de rentabilidad $\{r(t),\ t > 0\}$ es autosemejante de exponente ½:

$$r(n\tau) \cong n^{1/2} r(\tau) \, ,$$

donde $\cong$ simboliza igualdad en distribución. En consecuencia, *el modelo estándar de las rentabilidades bursátiles que utiliza la teoría financiera es un modelo fractal.*

El segundo aspecto a subrayar es que el riesgo es un riesgo unidimensional, valorado únicamente en términos de la variación endógena de la rentabilidad del activo, la desviación típica o *volatilidad*. De aquí se sigue la importancia de la ley marginal del proceso que modela la evolución de las tasas de rentabilidades mediante la medida de la variación, y por ello del riesgo, de los activos financieros. La utilización de la distribución normal calibra el nivel del riesgo del mercado y por ello las posibilidades de dispersión de las rentabilidades efectivamente observadas en torno a su media teórica. La modelación gaussiana implica que el número de valores extremos observados sobre un mercado dado es muy pequeño en relación al número de valores próximos a la media, dicho de otro modo que la probabilidad de saltos de las cotizaciones o de rupturas de mercado, es extremadamente débil. Esta es la consecuencia fundamental de la hipótesis de normalidad de las rentabilidades.

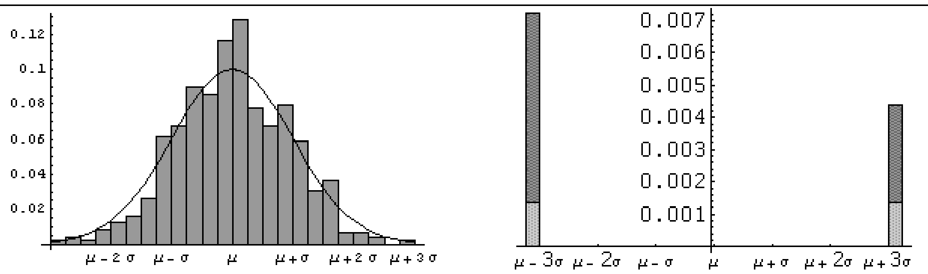


Figura 2: Distribución empírica de los incrementos diarios de la tasa de rentabilidad del Ibex35 durante la década de los 90 y sus colas acumuladas frente a una normal con la misma media y varianza.

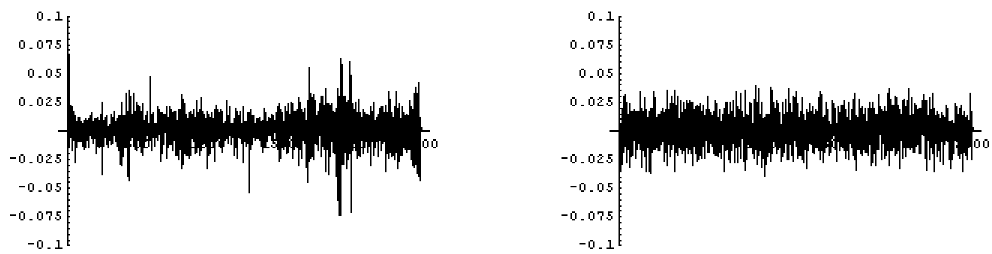


Figura 3: Incrementos diarios de la tasa de rentabilidad del Ibex35 durante la década de los 90 frente a los incrementos de un movimiento browniano correspondiente a una normal con la misma media y varianza.

La observación del comportamiento real de los mercados y, en particular, de las serie de precios financieros, revelan la ocurrencia de saltos, de discontinuidades, de rupturas mucho más frecuentes que las predichas por la ley normal: si se observa la forma de las distribuciones empíricas de las rentabilidades empíricas que se dan en los mercados, se verifica que éstas presentan un carácter de leptocurtosis, es decir, poseen colas espesas, contienen un gran número de valores elevados en relación a su centro. Dicho de otro modo, las dispersiones potenciales de las cotizaciones son más fuertes que las previstas por la modelación estándar: la normalidad de las rentabilidades no se verifica empíricamente. El proceso de evolución de las rentabilidades es más *quebrada e irregular* que un proceso de incrementos estacionarios e independientes gaussiano, que no presenta saltos.

Por tanto, parece necesario poner en cuestión al menos una de las tres hipótesis de la modelación: estacionariedad, independencia y distribución gaussiana de los incrementos. De antemano todas las vías son igualmente válidas, nos concentraremos, al menos de momento, en la hipótesis gaussiana. Con más precisión nos proponemos mantener la hipótesis fractal de la modelación financiera estándar, pero generalizando, siempre en el marco i.i.d., a los casos en que el exponente de autosemejanza valga $1/\alpha$, con α entre 0 y 2. Las distribuciones correspondientes, llamadas alfa estables, tienen como características principales tener una varianza infinita y conducir a procesos cuyas trayectorias son discontinuas. Estas distribuciones se adaptan a la modelación de fenómenos *más erráticos* que el movimiento browniano.

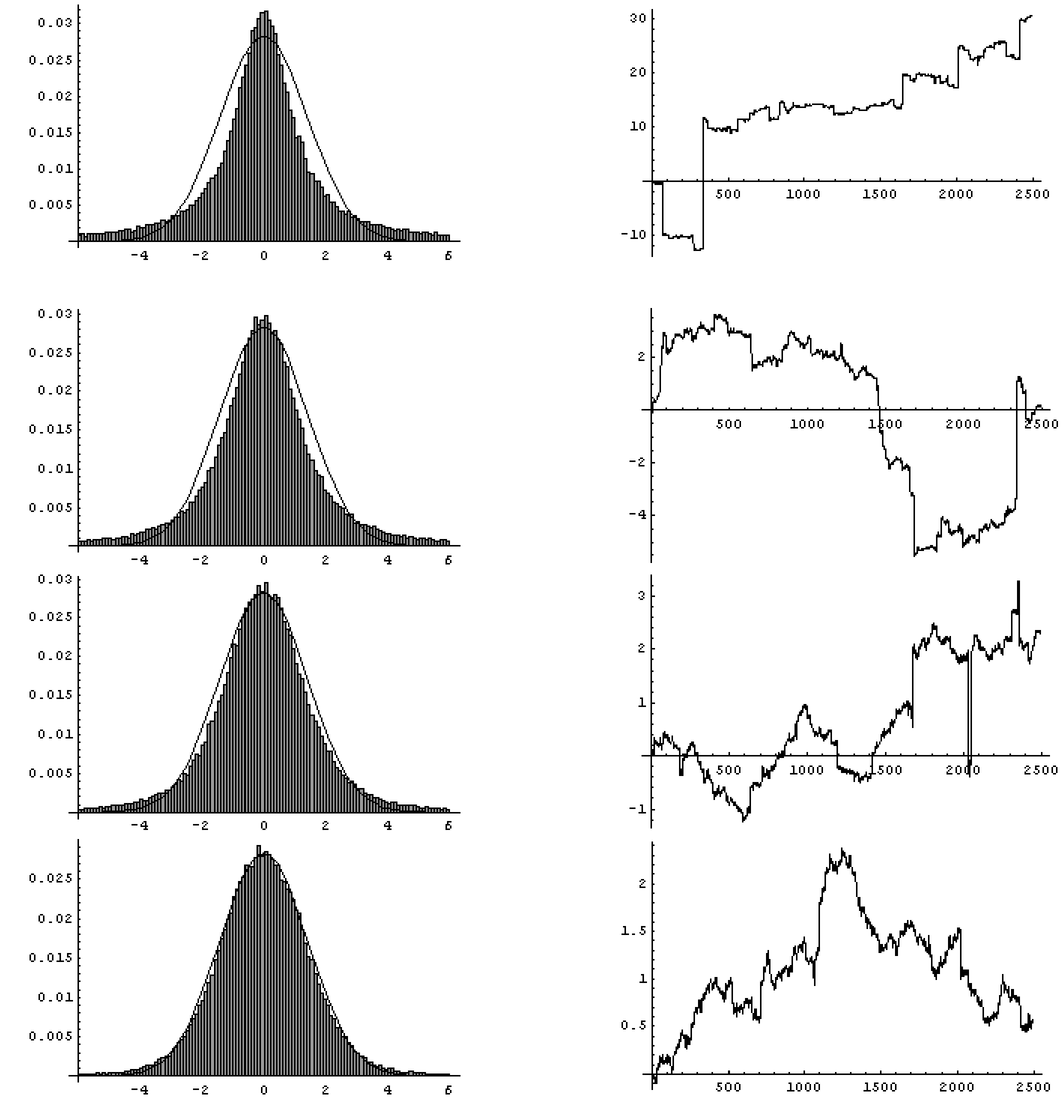


Figura 4: Distribuciones estables con α=1, α=1.25, α=1.5 y α=1.75 frente a una normal con los mismos parámetros de localización y escala junto a la simulación del proceso estable correspondiente.

Para situar el punto de partida del problema que nos proponemos, consideremos la siguiente propiedad de invarianza de escala de las rentabilidades bursátiles: según la modelación usual sobre una duración τ , las rentabilidades periódicas $\Delta r_{t,\tau}$ varían en la forma:

$$\sigma\sqrt{\tau}\ ,$$

donde σ es una constante positiva.

Para dar cuenta de la no normalidad observada sobre las distribuciones reales de las rentabilidades bursátiles, la investigación ha introducido modelaciones en las que $\Delta r_{t,\tau}$ varía en la forma

$$\sigma(t)\sqrt{\tau}\ ,$$

es decir, haciendo variar la volatilidad.

En relación a este enfoque, se propone mantener el factor σ ajustando el valor del exponente de τ :

$$\sigma\tau^{1/\alpha},$$

donde $1/\alpha$ es el exponente de autosemejanza.

Por una parte esta vía parece que puede ser prometedora para sus aplicaciones en la práctica financiera, por lo que se refiere a la relación entre complejidad de la modelación y ganancia inmediata de realismo. Por otra parte, esta generalización fractal sencilla del movimiento browniano dispone de una teoría matemática más desarrollada en los momentos actuales.

Bajo hipótesis i.i.d. consideraremos una generalización fractal del marco probabilista de la modelación financiera que tiene en cuenta tres parámetros, dos de ellos de dimensión del riesgo, que describen

1. La rentabilidad anticipada que es un parámetro de *localización* correspondiente a la media gaussiana μ;
2. La dispersión de las rentabilidades esperadas, que es un parámetro de *talla o escala*, que se corresponde con la desviación típica σ;
3. El espesor de las colas de la distribución, que cuantifica el peso de los sucesos raros, y por ello el número y la amplitud de los saltos posibles, es un parámetro de *forma* (en el sentido de la estructura de las variaciones) de la distribución que se corresponde con el exponente característico α .

De este modo se mantiene el concepto de eficiencia de los mercados que se hace más extenso con la introducción del tercer parámetro que es de forma del comportamiento errático del mercado. Los mercados fractales siguen siendo eficientes.

El concepto de eficiencia y las leyes de probabilidad.

El concepto de eficiencia informativa de los mercados se encuentra en el núcleo de la teoría financiera moderna y constituye su fundamento. Concepto de contornos fluctuantes y fronteras difusas. La intuición del mismo comienza con Bachelier (1900), y el esbozo en Cowles (1933). Aparece por primera vez en lo que se puede decir su formulación moderna en Samuelson (1965), donde define explícitamente el contenido del concepto así como su relación con las formulaciones probabilistas. En otras palabras, en un mercado eficiente informativamente, las variaciones futuras de las cotizaciones son imprevisibles si las anticipaciones de esas cotizaciones se hacen incluyendo la información de todos los que intervienen en el mercado. En ese caso, las variaciones de las cotizaciones evolucionan *al azar*.

Los aspectos informativos del concepto de eficiencia han sido ampliamente analizados por Grossman, S, y Stiglitz, J.E. (1980). Hoy es bien conocido que el concepto es en parte contradictorio. Para una buena síntesis de esas discusiones se puede ver en Roger (1988). Nos centraremos sobre las relaciones específicas entre eficiencia y leyes de probabilidad.

En cuanto al *impacto intelectual de las herramientas probabilistas* que operan en el interior del concepto, herramientas que no han sido objeto de un análisis crítico específico, en tanto que marcaban *maneras de pensar* la teoría de la eficiencia, hay algunos trabajos que han abordado este aspecto:

1. Trabajos sobre la relación entre eficiencia informativa y modelos de martingala, Mandelbrot, B. (1971).
2. Trabajos que adoptan un punto de vista más próximo a la historia de las ciencias: Longin, F. y Walter, C. (1994); Walter, C. (1996).

Estos trabajos de investigación epistemológica son importantes. En efecto, la eficiencia de los mercados es un concepto que hunde sus raíces en una larga historia, la del análisis estadístico de las variaciones bursátiles: los orígenes intelectuales del concepto de eficiencia son orígenes probabilistas y estadísticos, y la eficiencia de los mercados ha estado estrechamente asociada, desde sus comienzos, a un paradigma probabilista-estadístico. Las incomprensiones que resultan de un mal uso del concepto por parte de los profesionales del mercado son también la consecuencia de la confusión entre eficiencia y ley de probabilidad. Comprender la naturaleza de la articulación entre eficiencia informativa y modelo probabilista permite por ello aprehender mejor los problemas que se proponen hoy sobre los mercados bursátiles, y las razones por las que se podría conservar la hipótesis de eficiencia modificando las leyes de probabilidad que operan en ella. Es por esta razón por lo que se desarrolla las interacciones entre eficiencia y aleatoriedad.

Las definiciones clásicas de la eficiencia de los mercados

La definición de Fama (1970, pág. 383) es como sigue:

> *"Un mercado bursátil se dice <u>eficiente</u> si las cotizaciones registradas <u>reflejan plenamente</u> toda la información disponible".*

El subrayado de *<u>reflejan plenamente</u>* parece indicar tanto su importancia como la necesidad de precisar esa expresión.

La definición de Mandelbrot (1971, pag. 225) dice:

> *"Un mercado bursátil se dice <u>eficiente</u> si las cotizaciones registradas reflejan plenamente toda la información pertinente y disponible. La llegada de una nueva información al mercado produce una ocasión de arbitraje (una imperfección de información), pero se supone que toda imperfección es inmediatamente arbitrada, de manera que la sucesión de los precios arbitrados debe ser una martingala".*

Mandelbrot precisa la definición de Fama añadiendo a la información "disponible" la noción de información pertinente, es decir, utilizada realmente por el mercado para la fijación de los precios.

La definición de Jensen (1978, pág. 96) presenta el concepto de eficiencia (eficacia) bajo una forma operativa que permite su validación mediante comprobaciones empíricas:

"Un mercado bursátil se dice eficiente, en relación a un conjunto de información dado, si es imposible realizar una ganancia bursátil poniendo en práctica una política de intervención o de gestión fundamentada sobre este conjunto de información".

Lo que quiere decir que no se puede sacar partido de una ventaja informativa pues, como dicen los operadores de los mercados, la información útil está ya *en las cotizaciones*. Esta definición conduce a la introducción de eficiencia en relación a diferentes subconjuntos de información, y a distinguirla de la eficiencia en relación a una información global para todo el mercado.

En este sentido, Beaver (1981, pág. 35) propone la siguiente definición:

Un mercado bursátil se dice eficiente en relación a un conjunto de información dado, si la comunicación a todos los participantes del mercado de la información contenida en este conjunto específico no tiene impacto sobre el nivel de las cotizaciones registradas.

A partir de esta definición surge la aproximación moderna de la eficiencia, que modela la agregación de informaciones diferenciadas en la formación del precio de equilibrio, problemática introducida y desarrollada en Grossman (1976) y Grossman y Stiglitz (1980).

Por último, todos estos enfoques están reunidos en la definición convencional dada por Malkiel (1989):

Un mercado se dice eficiente si las cotizaciones registradas reflejan plena y correctamente toda la información disponible y pertinente. Formalmente, un mercado se dice eficiente en relación a un conjunto de información dado, si la comunicación a todos los participantes del mercado de la información contenida en este conjunto no tiene impacto sobre el nivel de las cotizaciones registradas. Además, la eficiencia de un mercado en relación a un conjunto de información dado tiene una consecuencia práctica: es imposible realizar ganancias bursátiles poniendo en práctica una política de intervención basada en este conjunto de información.

Esta definición de Malkiel representa la concepción que se puede denominar clásica de la eficiencia, en la medida en que no hace referencia a la problemática de agregación de información de los diferentes agentes. Es este concepto, en su forma clásica, el que ha pasado de la investigación a la industria y que ha tenido consecuencias profesionales importantes, como se verá más adelante.

Como última nota después de estas definiciones, es necesario resaltar que el término *eficiencia*, traducción usual de la palabra inglesa *efficiency*, no refleja bien la analogía física contenida en el concepto. Utilizando la palabra *eficacia* se ve que se cualifica mejor un mercado por su *eficacia* a la hora de transformar información en dinero. De la misma manera que la *eficacia* de un motor es su capacidad para transformar la energía contenida en el combustible en trabajo útil, la eficacia de un mercado se medirá por su capacidad para transformar sin pérdida toda la información en precios.

Por ello, se puede proponer la siguiente definición:

Un mercado bursátil se dice eficaz si transforma correctamente la información en precios.

Desde este punto de vista un mercado puede ser más o menos eficaz, y puede existir una pérdida (ruido) en ese proceso de transformación. El concepto de eficacia perfecta puede ser utilizado como patrón para apreciar la cualidad del instrumento *mercado* para transformar la información y comparar los mercados entre ellos según este criterio.

Aunque en adelante seguiremos utilizando la traducción financiera española de la palabra (efficiency), nos parece adecuado mantener en mente este concepto de *eficacia*.

Implicaciones profesionales de la restricción gaussiana

Las consecuencias operativas de la identificación de la eficiencia informativa con la distribución normal fueron inmensas. Un ejemplo bastante emblemático es el de la industria y empresas de gestión de activos, esta industria es representativa de la manera en que *"la investigación universitaria en finanzas ha modificado las prácticas profesionales"* Fama (1991, pág 1608). Modificación que resultó de la implantación del concepto de eficiencia en las cadenas de producción y de gestión de activos en EE.UU. primero, y después en Europa y Japón.

La definición de Jensen (1978) indicaba claramente que toda gestión activa conducida por gestores profesionales experimentados, y ayudados por analistas financieros que buscaban identificar los valores a comprar, todos esos esfuerzos no podían permitir realizar resultados superiores a los obtenidos por un índice representativo del mercado. Esa era la conclusión a la que había llegado Cowles (1933), para quien la mejor estrategia de inversión posible era la constitución de un portafolio (cartera) de mercado. Formalmente, eso significa que la esperanza matemática de la desviación de ganancias en relación al índice, resultante de una estrategia de inversión basada en el conjunto de información Φ_t, es nula para un nivel de riesgo idéntico. Esta manera de ver la eficiencia constituye el *fundamento conceptual de la gestión por índice llamada pasiva.* La eficiencia de los mercados aparece como *la última razón* de la gestión por índice pasiva.

Esta manera de ver la eficiencia remite, entre otras cosas, a una manera de *definir un tipo de aleatoriedad*. En efecto, si, en *media*, la desviación del resultado entre una gestión activa y un índice de mercado es nula, pueden existir sin embargo desviaciones temporales más o menos importantes según el tipo de aleatoriedad. Con una aleatoriedad gaussiana, las desviaciones estarán limitadas al *ruido* en torno a cero, y los partidarios de la gestión por índice podrán hacer valer mecánicamente un índice de mercado antes que comprometerse en costos importantes para el mantenimiento de equipos de gestores que, en definitiva, se revelan inútiles y que, según la célebre boutade de Samuelson, deberían reciclarse como cuadros en la industria. El reciclaje de Samuelson no es en realidad más que la consecuencia de una elección de la aleatoriedad gaussiana para la formalización del resultado. Con una aleatoriedad no gaussiana, la existencia de desviaciones importantes no está excluida, y una esperanza de desviación nula no está en contradicción con una gestión activa.

En efecto, sea X_n la desviación del resultado observado en la fecha *n* para el período $[n-1, n]$ entre un índice de mercado y una cartera gestionada activamente y del mismo riesgo. Se observa a lo largo del tiempo una sucesión de desviaciones que se suponen i.i.d. $\{X_n, n > 0\}$. El problema planteado por Cowles y Samuelson se puede analizar formalmente como sigue. Dadas una sucesión $X_1, X_2, ..., X_n$ de variables aleatorias i.i.d., se plantea la cuestión de la forma de su distribución límite de su suma $X_1 + X_2 + ... + X_n$ cuando $n \to \infty$. La solución a este problema fue dada por Paul Levy en 1937, con el descubrimiento de las leyes estables: cuando $E(X_i^2)$ es finita, la única ley límite es la ley de Laplace-Gauss. En los demás casos, nos encontramos en presencia de leyes más complejas que generalizan la ley normal (la ley de Cauchy es un ejemplo). En estos casos, la convergencia eventual de las medias empíricas no responde a los mismos comportamientos que en el caso gaussiano.

Por ejemplo, supongamos un caso extremo en el cual las desviaciones de los resultados siguen una ley de Cauchy, definida por

$$P\left[a < X < b\right] = \int_a^b \frac{dx}{\pi(1+x^2)} = \frac{1}{\pi}(arctg\ b - arctg\ a)$$

Para esta distribución, la esperanza matemática (la desviación del resultado medio) no existe, ya que la integral $\int_{-\infty}^{+\infty} \frac{xdx}{\pi(1+x^2)}$ no está definida. Las medias empíricas obtenidas por lanzamientos aleatorios de Cauchy no convergen y obedecen a comportamientos desconcertantes y contraintuitivos: la no existencia de la esperanza teórica tiene como consecuencia la aparición en las trayectorias de períodos de cuasi estabilidad, seguidas de saltos, después, de nuevo, de períodos de calma,… Todo ocurre como si, de repente, de manera inesperada las medias empíricas diesen un salto, en un sentido u otro, tanto más imprevisible como contrario a la intuición. Nos encontramos entonces delante de una configuración probabilista en la cual podrían existir durante largos períodos una desviación del resultado importante entre un índice de mercado y una gestión activa del mismo nivel de riesgo. En este caso una gestión por índice pasiva no sería la mejor política de gestión. La razón es que si algunos momentos no existen, la aleatoriedad de la rentabilidad puede variar de manera brutal. Para detalles sobre la historia de la creación de la gestión por índice que sigue la hipótesis gaussiana se puede ver Walter (2000). Por todo ello, se puede decir que la gestión por índice pasiva es una consecuencia, no de la hipótesis de eficiencia, sino de la hipótesis gaussiana.

Otra consecuencia de la reducción de la eficiencia a la hipótesis gaussiana es la estructuración de los procesos de inversión de las sociedades de gestión según la forma llamada en la jerga de la industria de gestión *approach top down*. Es el desarrollo del ahorro a largo plazo que corresponde a la constitución de los fondos de pensiones que ha conducido al desarrollo del actuariado financiero y a una racionalización de los procesos de inversión en la gestión de este ahorro. Se han desarrollado métodos de determinación de asignaciones llamadas *estratégicas* de activos o composición de carteras a largo plazo, en función de los objetivos de gestión de los administradores de los fondos de pensiones, métodos que se apoyan sobre las previsiones de estabilidad de las rentabilidades a largo plazo y la adecuación activo pasivo.

Los procesos de inversión se han construido privilegiando las asignaciones estratégicas de activos sobre las intervenciones llamadas *tácticas*, decisiones a corto plazo, y las elecciones de títulos, consideradas como de ruido de actuación alrededor de la base constituida por la asignación estratégica. En sus presentaciones comerciales han hablado del *triángulo de actuación* según el cual el 70% del resultado de una actuación proviene de la asignación estratégica, el 20% de las asignaciones tácticas, y solamente el 10% de las elecciones de títulos.

El principio subyacente a esta estructuración de los procesos de inversión es la idea según la cual *la paciencia reduce el riesgo*, Hammer (1991, pág. 48), y que por ello el ruido desaparece sobre períodos largo, haciendo converger los resultados reales hacia el resultado teórico de la asignación estratégica. Esta descomposición del resultado en fuentes que contribuyen al resultado total proviene de la aplicación del CAPM gaussiano. En efecto, en virtud de la aplicación del modelo de evaluación de los activos financieros de Sharpe-Lintner-Mossin, la rentabilidad esperada, y por tanto el resultado de una cartera se descompone en la parte que proviene del mercado y la parte que proviene de las elecciones de títulos. Se sabe que el riesgo específico (la elección de títulos) no está remunerado, la aleatoriedad gaussiana que garantiza un decrecimiento rápido de la varianza total.

Pero según el tipo de aleatoriedad, estable o gaussiana, será necesario ser más o menos paciente... El paradigma de la asignación estratégica y de la diversificación depende del tipo de aleatoriedad, y se puede por tanto adelantar que el postulado de la asignación estratégica es una idea que proviene, no de la hipótesis de eficiencia, sino de la hipótesis gaussiana.

Las anomalías observadas en relación a la restricción gaussiana

Los trabajos sobre la comprobación empírica de la hipótesis gaussiana son bastante numerosos. Existen excelentes síntesis de las que cabe citar Taylor (1986), Campbell, J. Lo A. y MacKinlay (1997). Los dos grandes ejes que emergen de estos trabajos hacen hincapié en dos tipos de fenómenos calificados de anomalías sobre los mercados bursátiles:

1. Una aparente previsibilidad de las rentabilidades, que ponen en evidencia dependencias o correlaciones sobre diferentes horizontes de tiempo.
2. Una leptocurtosis (no normalidad de las distribuciones empíricas) que parece atenuarse con el paso del tiempo, fuerte en alta frecuencia, más débil en la baja frecuencia.

Estas dos formas de anomalía pueden ser conectadas entre ellas en la medida que se puede resolver una parte del problema, el de la leptocurtosis de las distribuciones marginales, introduciendo dependencias no lineales sobre los parámetros de las distribuciones condicionales.

Esto está en el origen de los modelos ARCH que orienta la investigación hacia la dependencia de las volatilidades más que de las rentabilidades. Se ha demostrado cómo formas de dependencia no lineales pueden producir distribuciones marginales no normales.

Los trabajos que se citan a continuación, Baillie y King (1996), Lordic y Mignon (1999), son buenas síntesis sobre la investigación de la previsibilidad (dependencias lineales y no lineales) en los mercados. Nos centraremos solamente en el problema de la no normalidad.

Los primero trabajos donde se da cuenta de que las colas de las distribuciones de las rentabilidades empíricas contenían demasiados puntos para que pudieran ser ajustadas por una ley Laplace-Gauss aparecen en Osborne (1959), Alexander (1961), Cootner (1962), pero consideran el fenómeno como despreciable.

El primero en atraer la atención sobre la leptocurtosis de las distribuciones observadas, así como de la necesidad de no despreciar e ignorar el fenómeno se encuentra en los trabajos de Mandelbrot (1962, 1963a). Según Mandelbrot, era necesario interesarse de forma particular en esos puntos, que eran el refugio de puntos aberrantes que contenían informaciones esenciales sobre el funcionamiento de los mercados, para la comprensión de riesgos potenciales. Según esta apreciación, *la información importante no se pierde en la suma* (las medias de las distribuciones), *sino que se refugia en las colas*. La posición de Mandelbrot significa una variación radical de la visión de las variaciones bursátiles. Hasta entonces, se habían interesado en las fluctuaciones de las sumas parciales de las variables aleatorias, y en el comportamiento asintótico de las medias empíricas. En Mandelbrot (1963b) sugería *interesarse por las fluctuaciones extremas.*

Mandelbrot (1962) utiliza por primera vez la ley de Pareto para el estudio del comportamiento del mercado del algodón desde 1880 hasta 1958, en el que observa que rentabilidades $\Delta_{t,1}$ y $\Delta_{t,20}$, donde la unidad de tiempo es el día, son demasiado estiradas para ser gaussianas. Aplica de nuevo este análisis en los mercados financieros, Mandelbrot (1963a), y obtiene resultados parecidos.

Fama (1965a), aplica este análisis al índice Dow Jones, y confirma el resultado de Mandelbrot.

Mandelbrot (1967), vuelve a aplicar el mismo análisis en otros mercados, con frecuencias semanales, mensuales y anuales. En particular, analiza las cotizaciones de trigo entre 1883 y 1936; cuatro acciones de sociedades de ferrocarriles, de 1857 a 1936; la tasa de cambio USD/STG, de 1803 a 1895. Los resultados de estos trabajos confirman la leptocurtosis de las distribuciones empíricas.

Muñoz San Miguel (2002, pág. 170), estudia la serie de cierres diarios del IBEX35 en la década de los noventa del siglo pasado y comprueba que la distribución empírica de los incrementos logarítmicos diarios, semanales y mensuales de la citada serie presentan un fuerte nivel de curtosis marcadamente positiva, las distribuciones están ligeramente sesgadas hacia la izquierda y las colas son sensiblemente mayores que las correspondientes a una distribución normal.

Aunque la no normalidad es considerada hoy día como un hecho básico de los mercados, es necesario subrayar que fueron necesarios casi 40 años para que ese hecho *observable* se convirtiera en un hecho *observado* y entrara en el campo de investigación de las finanzas. A pesar de las anomalías constatadas, las grandes variaciones bursátiles no tenían existencia intelectual en el campo de investigación de las finanzas. Así hasta la década de los 90 del siglo pasado estudiar los movimientos extremos comenzó a ser un enfoque original en las finanzas. Parece que ese ya no es el caso en nuestros días, donde la necesidad de tener en cuenta una estructura diferente del riesgo, y por ello de la aleatoriedad, en las distribuciones reales, está en el origen de una cierta convulsión en las formas de considerar los mercados.

En efecto, paralelamente a este cambio de visión en la comunidad científica, ha aparecido entre los profesionales una amplia toma de conciencia de la no normalidad y de sus consecuencias. Un seguimiento de la prensa profesional durante estos últimos años lo ilustra: entre 1998 y 2000, en una de las principales revistas profesionales no universitarias, *Risk,* aparecen mas de quince artículos consagrados al problema de la no normalidad. La leptocurticidad se presenta como uno de los diez principales problemas a abordar por la industria financiera para los próximos años[1], y la gestión de los sucesos raros se convierte en el problema clave a resolver en la gestión del riesgo[2]. Así por ejemplo, se puede leer en *Risk* que "la teoría de los valores extremos es una potente herramienta para la gestión de los riesgos"[3].

Este nuevo interés profesional procede de lo que, en el campo de la industria financiera se muestra, desde hace quince años, como un enfoque institucional de los riesgos del mercado,

---

[1] "The 10 challenges left to tackle", *Risk,* abril 2000, pp. 34-39.
[2] "Taylor-made for tails" *Risk,* febrero 2000, pp. 95-98.
[3] "History repeating", *Risk,* Enero 1998, pp. 99.

enfoque que se denomina prudente, y que se traduce en la existencia de grupos de reflexión internacionales cuyo objetivo es promover reglas de cuantificación de los riesgos universalmente aplicables por las entidades financieras en todos los dominios de su actividad. Esta prudente preocupación está relacionada con el aumento extremadamente rápido de los endosos de instrumentos financieros derivados, cuya inestabilidad y dificultad de gestión constituyen fuentes de enorme preocupación para las autoridades que tutelan el mercado: una muestra representativa de sesenta establecimientos financieros activos sobre los mercados de derivados, perteneciente al grupo de los diez países más ricos (G10), hace aparecer un endoso nominal sobre esos instrumentos de 130 billones de dólares. La exposición al riesgo de los establecimientos concernidos (relación entre los endosos nominales y los activos subyacentes) puede alcanzar el coeficiente 34, lo que quiere decir que el establecimiento está expuesto por 34 veces del valor de los activos de base sobre los cuales se negocian los productos derivados (fuente: Basle [1999], pág. 4).

Las instancias representativas que constituyen estos grupos de reflexión ven cada vez con más claridad que el fenómeno de la leptocurtosis debilitaba la fiabilidad de las herramientas y los instrumento basados en la distribución normal, y era necesario una renovación conceptual de los enfoques tradicionales.

Los procesos alfa estables es una de las maneras de modelar la aparición de discontinuidades en las series temporales. En efecto, las trayectorias de los movimientos alfa estables, que son los procesos fractales leptocúrticos que se utilizan actualmente para la modelación, presentan la propiedad de saltar de manera arbitrariamente grande entre dos realizaciones, y por ello presentar un gran número de puntos en las colas que corresponden a los valores extremos observados. La evidencia de distribuciones leptocúrticas se puede interpretar como un signo de la fractalidad de los mercados, en un contexto de independencia de los crecimientos.

Bibliografía

BACHELIER, L. (1900), “Théorie de la Spéculation”, Thèse de doctorat, Annales de l’Ecole Normale Supérieur, troisiéme série, pp. 21-86.

BAILLIE, R.; KING, M. (1996), “Fractional Differencing and long memory processes”, Special issue of *Journal of Econometrics,* vol. 73, nº 1.

Basle Committee on Banking Supervision (1999), *Trading and Derivatives Disclosures of Banks and Securities Firms.*

BEAVER, W. (1981), “Market Efficiency”, *The Accounting Review*, vol. 56, pp. 23-37.

CAMPBELL, J. ; LO A. ; MACKINLAY A.C. (1997), *The Econometrics of Financial Markets*, Princeton University Press.

COOTNER P.A. (1962), “Stock Prices: Random versus Systematic Changes”, *Industrial Management Review,* vol. 3, pp. 24-45.

COWLES, A.(1933), “Can Stock Markets Forecasters Forecasts?”, *Econometrica* 1, pp.309-324.

BERNSTEIN, P. (1992), *Capital Ideas. The improbable origin of modern Wall Street*, The Free Press, New Cork.

FAMA; E. (1965a), “The Behavior of Stock Market Prices”, *Journal of Business,* vol. 38, nº 1, pp. 34-195.

FAMA, E. (1970), "Efficient Capital Market: a Review of Theory and Empirical Work", *Journal of Finance*, vol. 25, pp. 383-417, and discussion pp. 418-423.

FAMA, E. (1991), "Efficient Capital Markets:II", *Journal of Finance*, vol. 46, pp. 1575-1617.

GROSSMAN, S. (1976), "On the efficiency of competitive stock markets where traders have diverse information", *Journal of Finance,* vol 31, pp. 573-595.

GROSSMAN, S, Y STIGLITZ, J.E. (1980), "On the impossibility of informationally efficiente markets". *American Economic Review"*, vol. 70, pp.393-408.

HAMMER,D. (1991), *Dynamic Asset Allocation*, New York, Wiley.

HAYEK, F.A.(1945), "The use of Knowledge in Society", *The American Economic Review,* vol. XXXV, September 1945, Number Four.

JENSEN, M. (1978), "Some anomalous evidence regarding market efficiency", *Journal of Financial Economics,* vol.6, pp. 95-101.

LONGIN, F. Y WALTER, C. (1994), "Scientific and Thecnological approach of knowledge: the case of financial markets" . Sixth Annual International Conference on Socio-Economics, HEC, juillet.

MALKIEL, B. (1989), "Is the Stock Market Efficient", *Science,* vol.243, pp. 1313-1318.

MANDELBROT, B.B. (1962), "Sur certains prix spéculatifs: faits empiriques et modèle base sur les processus stables non gaussiens de Paul Levy", *Compte-rendus à l'Academie des Sciences,* vol. 254, pp. 3968-3970.

MANDELBROT, B.B. (1963a), "The Variation of Certain Speculative Prices", *Journal of Business*, vol. 36, pp. 394-419.

MANDELBROT, B.B. (1963b), "New methods in statistical economics", *Journal of Political Economy,* vol. 71, pp. 421-440.

MANDELBROT, B.B. (1966), "Forecasts of future prices, unbiased markets, and martingale models", *Journal of Business,* vol. 39, pp. 242-255.

MANDELBROT, B.B. (1967), "The variation of Some Other Speculative Prices", *Journal of Business,* vol. 40, pp. 393-413.

MANDELBROT, B.B. (1971), "When can price be arbitraged efficiently?. A limit to the validity of random walk and martingale models", *Review of Economics and Statistics,* vol. 53, pp. 225-236.

MUÑOZ SAN MIGUEL, J. (2002), *La dimensión fractal en el mercado de capitales*. Tesis Doctoral. Universidad de Sevilla. Sevilla, Junio 2002.

OSBORNE, M.F.M. (1959). "Brownian Motion in the Stock Markets", *Operations Research,* vol. 7, pp. 145-173, and discussions: vol. 7, pp. 807-811.

ROGER, P. (1988), "Théorie des marches efficients et asymetrie d'information: une revue de la literature", *Finance*, vol. 9, nº 1, pp. 57-98.

SAMUELSON, P.A. (1965). "Proof that properly anticipated prices fluctuate randomly", *Industrial Management Review, vol. 6,* pp. 41-49.

TAYLOR, S. (1986), *Modelling Financial Time Series,* J.Wiley and Sons.

WALTER, C. (1996), "Une histoire du concept d'efficience sur les marchés financiers". *Annales. Histoire, Sciences Sociales*, vol. 51, nº 4, (jouillet-août), pp. 873-905.

WALTER, C. (2000), "The Efficient Market Hypothesis, the Gaussian Assumption, and the Investment Management Industry",

http://papers.ssrn.com/paper.taf?Abstract_id=267443

Capítulo 21

Karl Pearson, el creador de la estadística matemática

M. A. GÓMEZ VILLEGAS
Universidad Complutense de Madrid

Introducción

Karl Pearson fue historiador, escribió sobre folklore, fue un socialista convencido, abogado, matemático aplicado, biómetra, estadístico, maestro y biógrafo. Pero sin duda su contribución más importante es al nacimiento de la *Estadística Aplicada*. Es por lo que le debemos el mayor crédito, en frase de el mismo *"Hasta que los fenómenos de cualquier rama del conocimiento no hayan sido sometidos a medida y numero, no se puede decir que se trate de una ciencia"*.

Introdujo el *método de los momentos* para la obtención de estimadores, el *sistema de curvas de frecuencias* para disponer de distribuciones que pudieran aplicarse a los distintos fenómenos aleatorios, desarrollo la *correlación lineal* para aplicarla a la teoría de la herencia y de la evolución. Introdujo el *método de la* χ^2 para dar una medida del ajuste entre datos y distribuciones, para contrastar la homogeneidad entre varias muestras, y la independencia entre variables. Fundo los *Anales de Eugenesia* y en 1900, junto con Galton y Weldon, fundo la revista *Biometrika* de la que fue editor hasta su muerte. En una descripción autobiográfica decía *"una explicación para mi vida, se debe a una combinación de dos características que he heredado: capacidad para trabajar mucho y capacidad para relacionar las observaciones de los demás"*.

Datos biográficos

Nace en Londres en 1857 y muere en 1936, su familia es originaria de Yorkshire. Hijo de un abogado, estudia en el University College School. En 1873, a la edad de 16 años fue retirado de la escuela por motivos de salud, y pasa el año siguiente con un preceptor privado. En 1875 obtuvo una beca para el King's College, en Cambridge. El decía que Cambridge le dio, placer en las amistades, placer en las polémicas, placer en el estudio, placer en la búsqueda de nuevas luces, tanto en las matemáticas como en la filosofía y la religión; así como ayuda para mantener su radicalismo científico dentro de límites moderados y razonables. Con 22 años marcha a Alemania y estudia leyes, física y metafísica. Entre 1880 y 1884 es profesor de matemáticas en el King College y en el University College. En 1911 fue el primer profesor de Galton de Eugenesia, la naciente parte de la Biología encargada de los estudios encaminados a conseguir la mejora de las especies. Era un darwinista convencido.

En el año 1890 se producen dos sucesos importantes para la trayectoria científica de Pearson; Galton publica su *Herencia Natural* donde incluye sus trabajos sobre correlación y regresión y Weldon se incorpora a la cátedra de zoología en el University College de Londres. Los primeros trabajos le van a dotar de una herramienta, con la que cuantificar las medidas de dependencia con la que va a poder contrastar, con resultado positivo, la teoría de la evolución introducida por Darwin. La figura de Weldon le va a permitir trabajar con un biólogo que compartía sus ideas de la evolución y que sería una fuente inagotable de cuestiones, que obligarían a Pearson a ir obteniendo técnicas estadísticas que le permitieran responder a los problemas que Weldon le planteaba. Entre 1891 y 1892 imparte conferencias sobre la geometría de la estadística en el Gresham College, y en ellas introduce los estigmogramas, entigramas, histogramas, cartogramas, stereogramas, etc. Estas lecturas marcan el comienzo de una nueva época en la teoría y en la práctica de la estadística.

Entre 1893 y 1906 publica unos 100 artículos sobre la teoría estadística y sus aplicaciones. La capacidad de investigación de Pearson es asombrosa, a lo largo de su vida publico mas de 650 artículos, fundo junto con Galton y Weldon, en 1901, la revista *Biometrika* para publicar artículos de estadística aplicada a la biología, ese mismo año publica sus *Tablas para Estadísticos y Biometristas* para ayudar en los ajustes de curvas. En 1905 publica el artículo *Sobre la teoría general de la correlación asimétrica y la regresión no lineal*. En 1914 Fisher empieza la polémica con el cuando trata de publicar un artículo en *Biometrika*, sobre el coeficiente de correlación muestral para muestras de una población normal bivariante. El artículo fue referenciado por Weldon como biólogo y por K. Pearson como estadístico y fue rechazado. Posteriormente Fisher diría que su artículo había sido referenciado por un biólogo que no sabía estadística y por un estadístico que no sabía biología.

Para completar la personalidad de K. Pearson, decir que en su primera época, cuando descubre que los valores de la ruleta no son aleatorios, escribe al gobierno francés para que cierre los casinos y dedique el dinero a la Academia de Ciencias, para que se funde un laboratorio de probabilidad, que aplique esta al problema de la evolución biológica.

Contribuciones de K. Pearson

La primera contribución de K. Pearson que me interesa citar, sobre todo en este contexto, es su serie de conferencias sobre la Historia de la Estadística que dio en el University College de Londres entre los años de 1921 y 1933. Las conferencias fueron recogidas por su hijo Egon Pearson, catedrático de Estadística en el University College, y aunque algunas personas no eran partidarias de su publicación sin ser revisadas, constituyen un valioso documento para la historia.

Para hacerse una idea del tipo de trabajo que entraña transcribimos la siguiente cita de la introducción de las conferencias, tomada del prefacio de las conferencias dadas por K. Pearson.

> *Lleva mucho tiempo leer las fuentes originales. En la historia de la estadística muy poca gente se ha tomado la molestia de hacerlo. Yo podría dar muchos ejemplos, de la cantidad de errores que ha propiciado esta conducta, pero me contentare con poner tres o cuatro.*

La segunda contribución es la familia de curvas de K. Pearson. La siguiente contribución fue el método de la distancia de la χ^2 para dar una medida del ajuste entre una distribución teórica y una experimental.

El cuarto procedimiento que nos legó Pearson, fue la concreción de la definición del coeficiente de correlación lineal para el estudio de la dependencia estadística y el método de los momentos para determinar los parámetros desconocidos de una distribución, cuando se dispone de una muestra aleatoria simple de la misma.

La familia de distribuciones asimétricas

K. Pearson introduce la familia de distribuciones asimétricas como una alternativa a la distribución normal, que había sido la protagonista ya desde el tiempo de Quetelet. Llega a la familia de distribuciones razonando sobre una mixtura de dos distribuciones normales y concluye que puede haber situaciones en las que los errores de las observaciones no sean normales y por lo tanto se consigan mejores ajustes a situaciones prácticas mediante las mixturas. Los problemas técnicos en los que se ve envuelto son de envergadura, para la determinación de los parámetros se ve forzado a resolver una ecuación de grado 9. Esto es lo que le llevo a Galton a dudar de la corrección del método. No obstante fue, la resolución del problema de la mixtura lo que le hizo abordar el problema de la obtención de distribuciones que permitieran sustituir a la normal para modelizar la incertidumbre.
Introduce la familia de distribuciones en su publicación K.Pearson (1985), mediante la solución de la ecuación diferencial

$$\frac{1}{y}\frac{dy}{dx}=\frac{-x}{c_1+c_2x+c_3x^2}$$

obtiene, para valores convenientes de las constantes, la distribución beta simétrica, la distribución beta asimétrica, la gamma y la normal.
Además para ajustar los parámetros introduce el método de los momentos.

El método de la distancia de la χ^2

Esta contenido en una memoria de 1900 y lo introduce para dar una medida del ajuste entre una distribución de probabilidad y una muestra.

La idea es, dada la muestra $(x_1,\ldots,\ x_n)$ y la distribución $f(x\,|\,\theta)$ construir el estadístico

$$\sum_{i=1}^{k}\frac{(Y_i - y_i)^2}{y_i}$$

que se distribuye χ^2_{k-1} , si la muestra proviene de la distribución. Donde se supone realizada una partición de k elementos en el recorrido de la distribución, con lo que los valores Y_i , las frecuencias observadas de los x_i en el elemento i de la partición, pueden suponerse con distribución multinomial, e y_i son las frecuencias observadas bajo la hipótesis de que la distribución de la muestra es $f(x|\theta)$.

El procedimiento sería generalizado a los problemas de homogeneidad y a las tablas de contingencia, por el propio K. Pearson y por sus discípulos, Edgeworth y Yule, hasta culminar en los trabajos posteriores de Fisher. Información relevante de esta evolución, puede verse en Stigler (1986), el desarrollo de los métodos puede verse en Gómez Villegas (2005).

El coeficiente de correlación lineal

La medida de la independencia entre dos variables ha tenido una larga historia y ha preocupado, básicamente por su utilidad práctica, a bastantes científicos. Es Galton, el que consigue concretar su definición, aunque todavía incorrecta, pero es K. Pearson el que en dos memorias consigue precisarlo. La primera titulada "Regresión, herencia y panmixia" es de 1896; la segunda escrita en colaboración con Filon "Sobre los errores probables de las frecuencias y su influencia en la selección aleatoria, la variación y la correlación" es de 1898.

En la primera memoria, está incluida con precisión la definición del coeficiente de correlación muestral cómo

$$r = \frac{S_{xy}}{S_x S_y}$$

con $S_{xy} = \frac{1}{n}\sum_1^n (x_i - \overline{x})(y_i - \overline{y})$, $S_x^2 = \frac{1}{n}\sum_1^n (x_i - \overline{x})^2$ y $S_y^2 = \frac{1}{n}\sum_1^n (y_i - \overline{y})^2$ y también incluye la distribución del coeficiente de correlación poblacional ρ en el caso de una distribución normal bivariante. Curiosamente aplica un razonamiento bayesiano para determinar la distribución del coeficiente de correlación poblacional.

En la diferenciación entre el coeficiente de correlación muestral y poblacional, afirma que r es el estimador más probable de ρ , en concreto enuncia sin demostrarlo, que el valor que maximiza la distribución de probabilidad final que ha obtenido para ρ es el coeficiente de correlación muestral, con lo que anticipa el método de estimación de la máxima verosimilitud que posteriormente desarrollará Fisher.

En el verano de 1933 renuncia a su cátedra y se retira, el University College de Londres divide su cátedra en tres; una de Eugenesia que fue desempeñada por Fisher, una de Estadística que fue desempeñada por Egon Pearson, el hijo de K. Pearson, y una de Biometría. Puede decirse que en ese momento ha sido creada la estadística aplicada cómo un procedimiento para tratar la incertidumbre y para ser aplicada a todas y cada una de las ciencias experimentales.

Un estudio más detallado de la vida y del trabajo de K. Pearson puede consultarse en E. Pearson (1938).

Agradecimientos

Este trabajo ha sido realizado en parte con ayudas del Ministerio de Educación y Ciencia proyecto MTM2005-05462 y de la Comunidad de Madrid-Universidad Complutense proyecto 910395.

Bibliografia

GOMEZ VILLEGAS, M.A. (2005) *Inferencia Estadística*, Madrid: Díaz de Santos.

PEARSON, E.S. (1938) *An Appreciation of Some Aspects of His Life and Work*, Cambridge: Cambridge University Press (existe una traducción de A.Eidlicz (1948) Pearson Creador de la Estadística Aplicada, Buenos Aires: Espasa-Calpe).

PEARSON, K. (1900) "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling". *Philosophical Magazine* 5 th series, 50, 157-175.

PEARSON, K. (1978) *The History of Statistics in the 17 th and 18 th Centuries*, Edited by E.S. Pearson. New York: MacMillan.

PEARSON, K. (1895) "Contributions to the mathematical theory of evolution, II: skew variation". *Philosophical Transactions of the Royal Society of London*, A, 186, 343-414.

PEARSON, K. (1896) "Contributions to the mathematical theory of evolution, III: regresion. heredity and panmixia". *Philosophical Transactions of the Royal Society of London*, A, 187, 253-318.

PEARSON, K. AND FILON, L.N.G. (1898) "Contributions to the mathematical theory of evolution, IV: on the probable errors of the frequency constants and on the influence of random selection on variation and correlation". *Philosophical Transactions of the Royal Society of London*, A, 191, 229-311.

STIGLER, S.M. (1986) *The History of Statistics: The Measurement of Uncertainty before 1900*, Cambridge: Belknap Harvard.

Capítulo 22

Evolución de las investigaciones en poblaciones finitas

Una perspectiva metodológica

Jesús Basulto Santos
Domingo Martín Martín
Universidad de Sevilla
Santiago Murgui Izquierdo
Universidad de Valencia

Introducción

Sobre la Historia de la Inferencia en Poblaciones Finitas vamos a interesarnos más en las innovaciones metodológicas que surgieron en el siglo XIX y parte del XX, hasta 1934, que en las numerosas encuestas que se realizaron a partir del XIX.

Dejando de lado los Censos, la primera innovación que surgirá en el siglo XIX serán las llamadas Monografías de Le Play. La imposibilidad de extender los resultados de las Monografías a las poblaciones de interés será una justificación más para que el Método Representativo de Kiaer sea aceptado por los líderes del Instituto Internacional de Estadística. Éste utilizará la estratificación en sus encuestas y recurrirá a los censos para seleccionar muestras representativas. Desde su propuesta en 1895 deberá competir con el despegue de los censos, de las estadísticas generales de los recientes Institutos de Estadística y de las Monografías, hasta que en la reunión del Instituto Internacional de Estadística, celebrada en Berlín en 1903, se proponga un comité para estudiar "La aplicación del Método Representativo". No será hasta 1926 que por fin Bowley y Jensen presenten sus investigaciones sobre dicho método.

Bowley (1906) reflexionará sobre el porqué de la falta de aplicación del cálculo de probabilidades en la recogida de datos por muestreo, a pesar de que las encuestas venían siendo utilizados desde hace 20 años. Buscando una medida para el error de las estimaciones, Bowley propondrá el uso de intervalos aleatorios, basados en los métodos bayesianos y el teorema central del límite. Esto último servirá para medir la precisión de las constantes a estimar en las poblaciones finitas por medio de muestras seleccionadas aleatoriamente. Bowley utilizará el diseño estratificado con afijación proporcional como una alternativa al simple para aumentar la precisión de las estimaciones. El diseño por conglomerados será también propuesto. Ahora bien, ante el sesgo que produce un diseño aleatorio simple cuando se usan los promedios de la variable de interés con conglomerados de distintos tamaños, Bowley (1926) desarrollará la teoría de los diseños intencionados usando muestras balanceadas o equilibradas como solución al problema.

Por último, Neyman(1934) introducirá la innovación de los intervalos de confianza, una nueva aproximación para medir la precisión de las estimaciones, que será su "Método Representativo", y que le permitirá generalizar el diseño estratificado de Bowley y desarrollar los diseños aleatorios por conglomerados.

En el período que va desde 1855, con la publicación de Le Play de su monografía "Los Trabajadores Europeos" hasta el artículo de Neyman en 1934, se han presentado un conjunto de innovaciones sobre cómo hacer inferencias en poblaciones finitas que marcarán la senda a seguir en el futuro.

Si ahora consideramos este conjunto de innovaciones y las unimos a las que se han producido desde 1934 hasta final del siglo XX, podemos ver que hoy tenemos un abanico de métodos, cuya presentación, muchas veces formal, está muy alejado de su origen, lo que ante una situación concreta nos lleva a preguntarnos ¿qué método debemos utilizar?

Una forma de agrupar las distintas aproximaciones metodológicas a la inferencia de poblaciones finitas es la que establece la siguiente distinción: 1) Censos e investigaciones exhaustivas, 2) Muestras Representativas, 3) Muestras Aleatorias, 4) Muestras Intencionadas y 5) Muestras generadas de forma espontánea. Estas agrupaciones nos servirán para relacionar las innovaciones con sus orígenes históricos y, también, para señalar cuándo una es más conveniente que otra,

A partir de aquí, el trabajo está organizado en las secciones siguientes: En la sección primera resumimos la aproximación por censos e investigaciones exhaustivas donde reflexionamos sobre su utilidad, relacionando la innovación de las monografías con los censos. Reflexionamos sobre las muestras representativas en la sección dos, donde recogemos el Método Representativo de Kiaer. Las muestras aleatorias son recogidas en la sección tres, donde incluimos las innovaciones de Bowley, finalizando con la respuesta de éste al trabajo de Neyman. En la sección cuatro hablamos de las llamadas muestras intencionadas, donde recordamos la aportación de Bowley y los modelos de superpoblación. En la última sección tratamos las muestras generadas de forma espontánea. Finalizamos el trabajo con algunas conclusiones.

1ª Aproximación: Censos e Investigaciones Exhaustivas

Para conocer los valores que una variable Y toma sobre las unidades de un universo U, la primera opción a considerar es la realización de un censo o estudio exhaustivo, cuyo resultado es de desear que conduzca al conjunto de datos $(y_1, y_2, \ldots, y_N)$. Sin embargo, frecuentemente surgen errores y limitaciones que impiden conseguir el objetivo deseado. En cuanto a los errores es habitual establecer la siguiente clasificación:

a-Errores del marco o directorio:

Subcobertura, por existencia de unidades de U no consideradas al diseñar el censo

Sobrecobertura, por observación de unidades que realmente no pertenecen a U

Observación de unidades repetidas

b-Errores por falta de respuesta:

Falta de localización de alguna unidad

Falta de contacto con alguna unidad

Rechazo a colaborar

Incapacidad de observar la variable de interés

c-Errores de medida:

Discrepancia entre los verdaderos valores que se obtendrían con la medición de la variable y el valor recogido.

En cuanto a las limitaciones que suelen presentarse cabe distinguir:

1- De naturaleza presupuestaria
2- Asociadas con la duración de la investigación
3- Derivadas de la confidencialidad de la información

Un ejemplo histórico de investigación exhaustiva son "Las Relaciones Topográficas", respuestas a cuestionarios elaborados bajo el mandato de Felipe II en los años setenta del siglo XVI con el fin de conocer una descripción particular de los pueblos pertenecientes a Castilla. Los cuestionarios fueron enviados por orden del rey a los gobernadores, corregidores y otras justicias. Se ordenaba que los alcaldes y regidores comisionaran a vecinos del pueblo "personas antiguas y discretas y curiosas" que conociesen las materias que se preguntaban para que respondieran el cuestionario. Este censo que buscaba recoger todos los aspectos de los pueblos fue un fracaso ya que al final sólo se pudieron recoger 653 cuestionarios, y esto a pesar de remitir en muchos casos un segundo cuestionario simplificado del primero. Una estadística de los pueblos que contestaron a uno o a los dos cuestionarios es recogido por F. J. Campos y Fernández (2000). La falta de jurisdicción de corregidores y gobernadores en algunas villas y la ausencia de personas competentes para hacer las relaciones fueron los motivos más frecuentes del fracaso de esta investigación. Un análisis pormenorizado de la parte de los cuestionarios que recogen información sobre pesca ha sido realizado por S. Basulto en 2006 (Tesina de Licenciatura de la Universidad de Huelva).

La primera alternativa a los Censos fueron las Monografías que con su estudio de casos permitió recoger información más detallada. Debemos a Le Play el método de la Monografía, lo que hoy correspondería al método del caso.

Le Play fue un pionero de la Sociología y un impulsor del método científico en las ciencias sociales en el siglo XIX francés (Garrígós Monerris, 2003). Veamos una aproximación de este método en palabras de Adolfo Álvarez Buylla (1850-1927), catedrático de Economía Política y Hacienda Pública de la Universidad de Oviedo, que en 1903 tuvo que abandonar su tierra para incorporarse en Madrid al Instituto de Reformas Sociales[1] (1903-1924).

"*Es la mayor copia de datos para llegar al más exacto conocimiento del estado de los trabajadores, y a buscar la mejor organización posible de la sociedad, afín de que cesen tantas injusticias como dominan hoy por desgracia en las relaciones entre los hombres, … consiste en una minuciosa descripción del estado económico (inventario de bienes, ingresos, gastos de toda clase), moral, religioso, artístico, de cultura, educación, etc.;*

[1]Antecedente al Instituto de Reformas Sociales, fue la creación en 1883 de la Comisión de Reformas Sociales que hasta 1890 elaboró un extenso cuestionario de 223 preguntas con el objetivo de conocer la realidad obrera ("la cuestión social").

descripción sacada del examen directo y personal del modo de existencia del sujeto a observación, seguido de las consideraciones que sugiere al autor la situación de las familias obreras, con ánimo de encontrar la que, constituida con toda pureza de costumbres y bajo la singularidad del padre, viniera a ser como el tipo de familia cristiana, en cuya restauración creía encontrar Le Play la felicidad humana”.[2]

Un estudio de la obra “Les Ouvriers Europeens”, segunda edición de 1855, puede verse en el artículo de P.F. Lazarsfeld (1961).

En la mayoría de los casos, la realidad observada tras el intento de efectuar un análisis exhaustivo conduce a plantearse la siguiente cuestión: Si después de realizar una investigación censal únicamente se accede a observar la variable de interés en n unidades de un universo de tamaño N ¿será válida la extrapolación de los resultados que se obtengan a todo el universo completo? La respuesta a esta pregunta es condicional, dependerá de las características comparadas que posean el colectivo completo y aquel para el que se han obtenido los datos.

2ª Aproximación: Muestras Representativas

Una muestra se dice que es representativa del universo U con respecto a una serie de variables X_1, X_2,…, X_m, si la estructura de los datos obtenidos al observar las citadas variables en la muestra es la misma que la correspondiente a todo el universo. En coherencia con esta definición, cuando una muestra sea representativa, cualquier resultado acerca de las variables de referencia que se establezca sobre la muestra, podrá ser directamente expandido al universo completo y su validez estará plenamente garantizada.

Las limitaciones que presenta el concepto de representatividad son inherentes a la propia definición. Para poder asegurar que una muestra es representativa sería necesario conocer la estructura de la población completa, algo que ya de partida se considera el objeto de la investigación.

Matizando el concepto de representatividad, en la práctica estadística es frecuente desconocer la estructura de los datos poblacionales asociados con una variable de interés Y, y sin embargo disponer de la estructura correspondiente a una variable auxiliar X (podría ser la misma variable Y referida a una ocasión anterior). En tales circunstancias, el hecho de disponer de una muestra representativa con respecto a X ¿asegura la representatividad con respecto a la variable Y? En términos categóricos la respuesta a esta pregunta es negativa. No obstante, en la medida en que exista cierta vinculación entre ambas variables, será posible concebir aproximaciones a la idea formal de representatividad.

Fue el estadístico Noruego A. N. Kiaer quien en la reunión del Instituto Internacional de Estadística de 1895 celebrada en Berne, propuso el método representativo como una alternativa al censo. En el método representativo, la muestra reflejará la población finita y esto se puede lograr bien por muestreo equilibrado a través de una selección intencionada de las unidades muestrales, o por muestreo aleatorio. Este debate fue llevado a cabo por Kiaer en las reuniones bianuales de ISI celebradas en 1895, 1897, 1901 y 1903, encontrando resistencia por parte de los Institutos de Estadística Oficiales, partidarios de la realización de censos y muy poco inclinados ante la innovación que propugnaba Kiaer.

---

[2]Fernando, Pérez, J.M. (2007): Antonio Flores de Lemus: años de formación universitaria. Correspondencia con Francisco Giner de los Ríos. Real Academia de Ciencias Morales y Políticas. Página 130.

Los esfuerzos de Kiaer dieron sus frutos en la sesión del ISI celebrada en Berlin en 1903, cuando se adoptó una resolución que recomendaba el uso del método representativo sujeto a la provisión de que en la publicación de los resultados se especificaran las condiciones bajo las cuales se había realizado la selección de las observaciones. Veamos algunos de los trabajos que Kiaer presentó en las reuniones del Instituto Internacional de Estadística.

En 1895, en la reunión del Instituto Internacional de Estadística en Berne, presentó un trabajo con el título "Observations et expériences concernant des dènombrements représentatifs", donde introduce el método Representativo. Veamos alguno de sus comentarios.

En el caso de querer estudiar la llamada "cuestión social", Kiaer propone cuál es para él la población de interés.

> "*Una cosa me ha llamado la atención, y es que las investigaciones detalladas concernientes a los ingresos, las viviendas y demás condiciones económicas o sociales que han sido hechas en relación con las clases obreras, no hayan sido extendidas de una manera análoga a todas las clases sociales de la sociedad. Me parece evidente que aún centrándose sólo en la cuestión obrera propiamente dicha, se debe comparar la situación económica, social, moral, etc., de los obreros con la de las clases medias y las clases ricas. En un país donde las clases superiores son muy ricas y las clases medias tienen un nivel de vida muy holgado, las pretensiones de las clases obreras relativas a sus salarios, a sus viviendas, etc., se miden según una escala distinta que la de un país (o una localidad) en que la mayoría de las personas pertenecientes a las clases superiores no son ricas y las clases medias se encuentran en apuros.*
>
> *De esta proposición, que me parece del todo clara, se sigue que para apreciar bien las condiciones de la clase obrera, será necesario también conocer, además de ésta, los elementos análogos de las otras clases. Pero es necesario dar un paso más y decir que, puesto que la sociedad no consiste sólo de la clase obrera, no se debe descuidar, en las investigaciones sociales, ninguna clase de la sociedad.*"

Kiaer usará dos encuestas realizadas en Noruega donde aplicará el método representativo, que también llama enumeración representativa. La primera de las encuestas, que comenzó a elaborarse en 1894, pretendía crear una especie de seguridad social que cubriese las incapacidades debidas a la edad. La otra encuesta, que había comenzado a elaborarse en 1890, buscaba calcular la distribución de la renta del cabeza de familia según edad, estado civil y ocupación.

Vamos a considerar sólo la primera encuesta, donde el método Representativo fue aplicado de la forma siguiente: con la información del último censo, los cuestionarios fueron asignados proporcionalmente a la población de las ciudades y los distritos en que fueron agrupados los municipios. En el caso de las ciudades se eligieron las cinco con mayor población, además de la capital, que fue estratificada en cuatro segmentos definidos por calles según sus poblaciones. En cuanto a los pueblos, fueron clasificados en 18 grupos, según el censo de 1891, y cada grupo fue de nuevo clasificado según la importancia de su actividad económica. Vemos que hasta aquí Kiaer está haciendo uso de la estratificación tan empleada en nuestras encuestas modernas. En una segunda etapa, en el caso de la capital, se eligieron calles en cada uno de los grupos de forma intencionada buscando que representaran a todas las calles del mismo grupo. En la última etapa, se pidió a los entrevistadores que debían visitar todo tipo de viviendas, de tal manera que cubriesen las distintas condiciones sociales y económicas. Una vez asignados los cuestionarios, Kiaer comparaba la muestra con el censo respecto del sexo, ocupación y localización. En aquellos casos donde ocurrían discrepancias de la muestra con el censo, debido a que, por ejemplo, los entrevistadores no habían seguido las normas, Kiaer quitaba o añadia cuestionarios hasta lograr una buena aproximación de la muestra al censo.

¿Qué ocurrió con esta encuesta de Kiaer? Gracias al reciente trabajo de Einar Lie (2002) sabemos que esta encuesta fue un fracaso a pesar de que el tamaño era de 80.000 personas. Un problema que surgió fue que la muestra no recogió bien el hecho que debían haber más incapacitados en las ciudades que en los pueblos, también que la propensión a tener incapacidades era mayor en los solteros que los casados, y que finalmente la incapacidad aumentaba con la edad. Aunque Kiaer conocía que su muestra no recogía bien la incidencia de la discapacidad, dio su informe sin modificar la muestra. Einar Lie explica este comportamiento de Kiaer como una consecuencia de las limitaciones del método Representativo, en el sentido de que las variables de control que hacían representativa la muestra aunque estaban bien relacionadas con la variable de interés "número de incapacitados", problemas de selección de las últimas unidades rompieron estas relaciones. Nuestra opinión sobre este fracaso fue que la variable de interés "número de incapacitados" era un suceso raro y así era difícil investigar por medio del método representativo. Si la muestra hubiera sido aleatoria, acompañada de estratificaciones y variables de control, la muestra que necesitaríamos sería enorme. El mayor crítico a la encuesta de Kiaer fue Hjorth, un actuario, que según sus cálculos aproximados concluyó que el tamaño de la muestra era demasiado pequeño, lo que justifica nuestra opinión. Podríamos decir que Kiaer se enfrento con lo que hoy llamamos "poblaciones raras", que necesita la metodología del Muestreo por Redes (Monroe G. Sirken, 2004).

Volviendo a la reunión de Instituto Internacional de Estadística, la reacción de los colegas a este primer trabajo de Kiaers fue oponerse al método Representativo. Fue el profesor G. von Mayr, de la Universidad de Munich, quien afirmó que:

> "*Veo con mucho peligro el punto de vista recogido en su trabajo. Creo que la muestra representativa puede tener algún valor, pero éste está restringido al terreno ya iluminado por un censo. No podemos sustituir por cálculo la observación real de los hechos. La muestra sólo nos proporciona estadísticas para las unidades actualmente observadas, pero no verdaderas estadísticas de toda la población.*"

Esta respuesta de Mayr debe contemplarse dentro de los éxitos logrados por el método censal que condujo a sus usuarios a rechazar cualquier otra innovación que pudiera hacerle la competencia.

Kiaer volverá a presentar otro trabajo en la reunión del Instituto Internacional de Estadística en 1897 en St. Petersburgo. Con el título de "Sur les méthodes représentatives ou typologiques appliquées à la statistique", Kiaer nos dice lo siguiente :

> "*Por investigación representativa quiero expresar una exploración parcial donde la observación se hace sobre un gran número de lugares dispersos distribuidos sobre toda la extensión del territorio, de tal manera que el conjunto de los lugares formen una miniatura del territorio total. Estos lugares no deben ser elegidos de forma arbitraria, sino según una agrupación racional basada sobre los resultados generales de las estadísticas; y los cuestionario usados deben ser asignados a dichos lugares, de tal manera que los resultados sean controlados, respectos de sus diferencias, con la ayuda del censo general*"

En la reunión de Berne, un responsable, M. Bodio, resumía el estado de la cuestión sobre las monografías y los censos, con estas palabras "La monografía estadística y la enumeración son dos maneras de investigación de los hechos sociales que se complementan. La enumeración, por si sólo no puede dar más que perfiles generales de los fenómenos, la silueta, por así decir, de la figura.". Ante el olvido del Método Representativo, Kiaer afirmaba lo siguiente.

> "*Añadiría a ellos la investigación parcial en general. Al permitir poner el análisis en todos sus detalles de la vida económica y moral del pueblo, da la sangre, carne, nervios o esqueleto construidos por la estadística general, y a la vez la enumeración completa las nociones aportadas por la monografía".*

Y añade una afirmación que a veces es difícil de explicar.

> "Es necesario notar que el valor científico de las investigaciones parciales dependen mucho más de su carácter representativo, que del número de datos."

Kiaer presentará otro trabajo en la reunión del Instituto Internacional de Estadística en 1901 en Budapest. Con el título de "Sur les méthodes représentatives ou typologiques", donde quiso distinguir su método representativo del método de los Tipos. Kiaer nos dice lo siguiente.

> "*Una investigación detallada de un cierto distrito o de un cierto barrio de una ciudad, no es una investigación representativa. Si el distrito o el barrio son áreas tipos, estos métodos serán de Tipos. Pero no se podría generalizar los resultados para todo el país o toda la ciudad. Pero si examináramos gran número de localidades diseminadas de una manera proporcional a lo largo de los diferentes distritos del país o de las distintas partes de la ciudad, entonces podríamos generalizar los resultados por medio del método representativo.*"

En esta reunión Kiaer recibirá una nueva crítica por parte de Bortkiewicz. Este profesor de estadística de la Universidad de Berlín aplicará una prueba de significación a la hipótesis nula de que la muestra representativa, de uno de los trabajos de Kiaer, no se diferencia de la población de donde ha sido construida, y con métodos debido a Poisson, concluye que la hipótesis no soporta los datos observados.

Kruskal y Mosteller (1980) señalan, con acierto, como una metodología basada en un modelo probabilístico no puede ser aplicada a una muestra representativa, que al ser ésta intencionada está lejos de ser una realización de dicho modelo aleatorio. Se necesita una nueva metodología que justifique el uso de las pruebas de significación.

Kiaer participará en la reunión de Berlín de 1903, donde sus aportaciones, recogidas en la publicación de 1905, no añadirán nada a lo ya conocido hasta ese momento .

En la reunión de Berlín, el Instituto Internacional de Estadística propondrá un comité para estudiar "La aplicación del Método Representativo en Estadística". El comité estará formado por los siguientes profesores: A. L. Bowley, Corrado Gini, Adolph Jensen, Lucien Martch, Verrijn Stuard y Franz Zized. No será hasta la reunión de 1926 en Roma donde se presentarán las memorias de Jensen y de Bowley.

En la segunda década del siglo XX el método representativo fue ampliamente utilizado. Inicialmente el método preferido de selección muestral fue la selección intencionada pero gradualmente la aleatorización seria un fuerte competidor del muestreo equilibrado para seleccionar la muestra. Según la resolución del ISI de 1926, se puede obtener una muestra lo suficientemente representativa de la población mediante dos procedimientos: A) Selección aleatoria: se selecciona un número de unidades de forma que la regla dominante es que las probabilidades de inclusión sean iguales. B) Selección intencionada: donde se selecciona un número de grupos de unidades (en terminología moderna “conglomerados”) de forma que casi reproduzcan las mismas características que la población objetivo, a través de unas variables de control. (Recogido de las recomendaciones del ISI, reproducidas por Yates (1946))

Una aplicación interesante del método de Kiaer fue a la extracción de una muestra a partir de un censo realizado en un momento t para su comparación con otros censos realizados en momentos t’ posteriores. Una ilustración puede verse en el trabajo « Une application de la méthode représentative aux matériaux du dernier recensement de la population italienne » (1 er décembre 1921). Bull. Int. Satisti. Inst. 1928, vol. 23, 198-215. En él se propone guardar una muestra del Censo de Italia para su comparación con otros censos sobre temas que en el futuro fueran de interés. Se utilizaron variables particulares para elegir la muestra representativa, buscando que los valores medios fueran casi los mismos en la población y la muestra. La principal conclusión fue que al tomar otras variables, no usadas en la selección de la muestra representativa, se notó que las diferencias encontradas fueron anormales. Este trabajo fue el que Neyman uso en su crítica del Muestreo Intencionado.

Las experiencias señaladas permiten comprobar la imposibilidad de conseguir muestras fielmente representativas, aunque en ocasiones sí pueden alcanzarse aproximaciones. Esto nos conduce a plantearnos las siguientes cuestiones: Si existen indicios de que una muestra es aproximadamente representativa, ¿será correcto afirmar que los resultados muestrales serán aproximadamente válidos para la población completa?, ¿cómo se puede medir el grado de representatividad de una muestra? Y finalmente ¿cómo medir el grado de aproximación de los resultados muestrales a los correspondientes a la población completa?

3ª Aproximación: Muestras Aleatorias

Un cálculo del grado de aproximación que se obtiene al elevar a la población completa los resultados obtenidos sobre los datos muestrales puede resolverse asociando una estructura estocástica sobre los datos. Una forma, aunque no la única, de generar tal estructura consiste en introducir mecanismos aleatorios a la hora de seleccionar las unidades muestrales. El procedimiento de selección se especifica a través de un diseño muestral y éste a su vez explicita el soporte de la estructura estocástica inducida. Veamos los primeros pasos que se dieron para la aplicación del muestreo probabilístico originalmente debido a Arthur Lyon Bowley, hasta llegar al artículo de Neyman de 1934.

Bowley introdujo inn ovaciones en el Muestreo de Poblaciones Finitas. Varias de estas innovaciones fueron: (a) aplicar el cálculo de probabilidades al muestreo de poblaciones finitas, (b) desarrollar el diseño aleatorio simple y el diseño estratificado proporcional y (c) elaborar un muestreo intencionado para la selección de conglomerados. Recogemos a continuación los trabajos de Bowley relacionados con el Muestreo de Poblaciones Finitas.

(1) El artículo “Address to the Economics Statistics Section of the British Association for the Advancement of Science”, York, 1906, publicado en el Journal of the Royal Statistical Society, vol. 69, pp. 540-568.

(2) El libro de “Elements of Statistics”. 3rd. Ed. King and Son. London. 1907.

(3) El libro “Livelihood and Poverty” de A.L. Bowley and A.R. Burnett-Hurst. Vol.

VII, Rotledge/Thoemmes Pres, 1915. Edición de 1997.

(4) La memoria “Measurement of the Precision Attained in Sampling”, Bulletin of the International Statistical Institute, 22, 1926, Suplement to Liv. 1, [6].[62]. Preceded by summaries in French and English.

(5) La respuesta que Bowley dio en la discusión del artículo de Jerzy Neyman, 1934, “On the Different Aspects of the Representative Method: the Method of Stratified Sampling and the Method of Purposive Selection”, publicado en la Royal Statistical Society, pp. 558-625. El texto de Bowley que nos interesa está en la página 609.

En el artículo (1), Bowley señala el desconocimiento que se tiene de cómo calcular el error de una estimación, a pesar de tener una teoría de cómo hacerlo. En apoyo de la teoría cita a Gauss, Laplace, Quetelet, Edgeword y Karl Pearson. A continuación habla sobre la curva normal y su deducción de la binomial, y nos dice que las fórmulas encontradas tienen dos bases, la que suponen hipótesis a priori y la que se justifica por medio de ajustes empíricos. Todas estas fórmulas están disponibles y son aplicadas, por ejemplo, a los métodos simples de muestreo que existen desde hace 20 años. A continuación Bowley propone un intervalo aleatorio que lo usa para medir la precisión (la exactitud, como dice Bowley) de la constante que desea estimar de carácter cuantitativo, apoyado en la aproximación normal. También propone un intervalo aleatorio en el caso de querer estimar la proporción de un atributo. Ilustra su propuesta anterior mediante una simulación de muestras de una población finita usando un diseño sistemático. Este ejemplo empírico le sirve para conocer la aproximación normal, y también para mostrar el buen funcionamiento de las fórmulas de los intervalos. A continuación Bowley propone aplicar las fórmulas propuestas a la recogida de estadísticas de salarios, empleo, etc., por medio de muestras aleatorias.

En su libro (2), capítulo VII, trata ejemplos de toma de muestras en las ciencias naturales, señalando el ahorro, de tiempo y coste y, además, nos recuerda que las muestras son necesarias en los casos donde las unidades seleccionadas son sometidas a modificaciones que las hacen inservibles. A continuación nos dice que “lo esencial de un examen por muestreo es que todo miembro del grupo considerado debe tener casi la misma chance de ser incluido en la muestra”. En este capítulo, Bowley vuelve al ejemplo descrito en su artículo (1).

El libro (3) escrito con A.R. Burnett-Hurst contiene varias investigaciones realizadas por muestreo aleatorio en Northampton, Warrington, Stanley y Reading sobre las condiciones sociales de la llamada clase obrera. El capítulo que nos interesa aquí es el VI de metodología.

En este capítulo VI del libro, Bowley señala los errores que pueden ocurrir en una investigación por muestreo, critica la falta de una medida de precisión de los resultados y así, cuando quiere calcular los errores de las estimaciones, ilustra el uso de los intervalos aleatorios propuestos en su artículo de 1906 por medio del siguiente ejemplo: nos dice que si en la muestra de hogares de Reading tenemos que el 20% de los hogares tiene cuatro habitaciones, entonces el porcentaje de hogares, de la población bajo estudio, con cuatro habitaciones es tan verosímil de estar en el intervalo [19,06-20,94] como de estar fuera de él.

La memoria de 1926 (4) es el trabajo teórico que Bowley aporta sobre inferencia en poblaciones finitas. La memoria, de 57 páginas, está dividida en: Selección Aleatoria, Selección Intencionada y Pruebas Generales. Usando la versión bayesiana del teorema central del límite debido a Edgeworth, Bowley estudiará los diseños aleatorios simple y estratificado con afijación proporcional y una teoría sobre muestreo intencionado

Veamos por último la respuesta de Bowley al trabajo de Neyman (5). Nos interesa la parte de la página 609 donde encontramos lo siguiente.

> *"Given that in a sample of 1000 taken at random, there are 1 in 10 with the defined quality, and given that the population from which the sample was drawn contained any proportion between 120 and 80 per thousand, then the chance of such an occurrence is less than one in twenty (approx.). Actual figures, of course, do not matter. That margin between 120 and 80 per thousand in the assumed population is shown on the vertical of the confidence belt in the very illuminating graphs which Dr. Neyman has given. Does that really take us any further? Do we know more than was known to Todhunter? Does it take us beyond Karl Pearson and Edgeworth?Does it really lead us towards what we need - the chance that in the universe which we sampling the proportion is within these certain limits? I think it does not. I think we are in the position of knowing that either an improbable event has occurred or the proportion in the population is within the limits. To balance these things we must make an estimate and form a judgment as to the likelihood of the proportion in the universe- the very thing that is supposed to be eliminated. I do not say that we are making crude judgments that everything is equal throughout the possible range, but I think we are making some assumption or we have not got any further. I do not know that I have expressed my thougths quite accurately, but it is not a thing that has occurred to me for the first time this evening; it is the difficulty I have felt since the method was first propounded."*

La anterior cita está en inglés porque no es fácil saber qué es lo que quiere decir Bowley. Veamos nuestra interpretación de esta cita.

En primer lugar, Bowley calcula un intervalo, [80,120], en tanto por mil, que sirve para estimar la proporción poblacional, en tanto por mil, a partir de una muestra de tamaño 1000 con una proporción observada del 10%. Este intervalo se ha construido con una probabilidad de 95%, la aproximación normal a la binomial. Tomando como error de estimación el valor $\sqrt{\frac{0,1(1-0,9)}{1000}} = 0,00948$, al multiplicarlo por el percentil 0,975, igual a 1,96, obtenemos el intervalo de confianza anterior. Bowley nos dice que este intervalo contiene a la proporción poblacional con una chance 0,95. (Bowley afirma que "la chance of such an occurrence is less than one in twenty", una afirmación confusa). Vemos que Bowley interpreta el intervalo como un intervalo fijo que contiene a la proporción poblacional con una chance del 0,95. En otras citas que hemos recogido anteriormente puede verse que Bowley interpreta el intervalo como un intervalo bayesiano.

A continuación Bowley se pregunta si ¿hace falta realmente algo nuevo como eso? (se refiere a la construcción del intervalo de confianza por Neyman), y añade ¿podemos conocer más de lo que fue conocido por Todhunter?, y, también, ¿necesitamos ir más allá de Karl Pearson y Edgeworth? Es decir, para Bowley el cálculo y la interpretación del intervalo propuesto por él no necesita más añadidos, tal como le enseñaron sus maestros. La respuesta a estas preguntas es clara "I think it does not", es decir, "pienso que no viene al caso", y lo justifica porque del intervalo [80, 120] podemos afirmar o que no contiene a la proporción, en tanto por mil, de la población, o que si la contiene; la primera alternativa, dice, es improbable. A continuación Bowley nos dice sobre las dos alternativas que "para equilibrar estas cosas debemos construir una estimación y hacer un juicio sobre la verosimilitud de la proporción en el universo", es decir, Bowley, al igual que hace en su Memoria de 1926, propone que calculemos la función de verosimilitud, viendo qué valores de la proporción poblacional son más o menos verosímiles.

Bowley finaliza esta parte de su respuesta a Neyman diciendo que "lo esencial de lo que es supuesto debe ser eliminado", es decir, Bowley rechaza la propuesta de Neyman. Y añade que "No digo que hagamos juicios aproximados cuando decimos que todo es igual en cualquier parte del recorrido", es decir, suponer una distribución a priori uniforme sobre el rango de la proporción poblacional; "pero pienso que o hacemos algún supuesto o nos quedamos con lo que tenemos".

Es claro que la nueva innovación que propone Neyman choca de frente con el método tradicional bayesiano de hacer inferencias, lo que causa en Bowley problemas al expresarse debido a que le resulta difícil entender el intervalo de Neyman.

El clásico artículo de Neyman de 1934 aportó los fundamentos teóricos para el enfoque del muestreo probabilístico en la inferencia con encuestas muestrales. Puso de manifiesto teóricamente y con ejemplos prácticos que en problemas de gran escala, el muestreo aleatorio estratificado es preferible al muestreo equilibrado, debido a que este último puede lograr pobres resultados si los supuestos de modelos subyacentes son violados. Como hemos recogido anteriormente, Bowley y sus discípulos utilizaron únicamente diseños muestrales con probabilidades de inclusión iguales para cada unidad de la población, justificando este método de muestreo en su teórica capacidad para generar muestras representativas del universo. Neyman (1934) rompe esta camisa de fuerza al proponer un muestreo estratificado con afijación "óptima" y muestreo de conglomerados con estimación tipo "ratio". En ambas situaciones, se obtienen estimaciones validas de los totales, medias y proporciones de la población sin confiar en una muestra representativa seleccionada a través de un diseño con probabilidades de inclusión iguales (Rao y Bellhouse, 1990).

El efecto inmediato del artículo de Neyman fue establecer la primacía del método de muestreo aleatorio estratificado sobre el método de selección intencionada. Además, puso las bases de la inferencia basada en el diseño del muestreo probabilístico.

Las ventajas del muestreo probabilístico, a saber:

- Mayor alcance.
- Coste reducido.
- Menor tiempo de ejecución.
- Posibilidad de realizar inferencia sin necesidad de especificar un modelo (es suficiente que la muestra sea grande para que pueda ser aplicable la aproximación a la curva Normal).

Todo ello indujo un incremento del numero y tipo de encuestas obtenidas por muestreo probabilístico. De esta forma el enfoque de Neyman fue casi universalmente aceptado por los estadísticos que trabajaban en las investigaciones de poblaciones finitas.

El muestreo intencionado o muestreo equilibrado fue relegado a un segundo plano y quedó para encuestas a pequeña escala y para determinadas áreas como las encuestas de mercado. Observemos que en el enfoque basado en el diseño aleatorio, los valores de la población son tratados como fijos y las inferencias se derivan de la distribución utilizada para seleccionar las unidades de la muestra, lo que requiere que la muestra sea probabilística, de lo contrario, no se pueden realizar inferencias válidas.

Según Brewer (2002), la teoría rigurosa de la inferencia basada en el diseño fue definitivamente establecida como enfoque dominante con los libros de texto de Yates (1949) y Deming (1950), e incluso más influyentes, aquellos que escribieron Hansen, Hurwitz y Madow (1953) y Cochran (1953). Fue posteriormente, con la publicación del libro de Särndal, Swensson y Wretman (1992), cuando el enfoque asistido por modelo fue definitivamente aceptado. Bajo este enfoque las inferencias siguen estando basadas en la distribución utilizada para seleccionar las unidades en la muestra, pero es posible dotar a la población de la estructura que induce un modelo de superpoblación y utilizarla para seleccionar estimadores más adecuados.

Rao (1996) define la inferencia basado en el diseño de la siguiente forma: "La teoría de encuestas tradicional utiliza el muestreo probabilístico para seleccionar la muestra y para la inferencia de los datos muestrales. Este enfoque, basado en el diseño, conduce a inferencias de muestreo repetido validas, independientemente de la estructura de la población, incluso en situaciones complejas, al menos para muestras grandes. A menudo se utilizan modelos de población plausibles en la etapa de diseño de una encuesta para ayudar a elegir buenos diseños y estimadores, pero las inferencias están libres de especificación de modelos excepto algunos débiles supuestos tal como que la población tiene el suficiente buen comportamiento como para satisfacer el teorema central del límite".

En el enfoque basado en el diseño la muestra no hace falta que sea representativa, la condición indispensable es que la muestra se seleccione aleatoriamente con un tamaño suficientemente grande. La "falta de representatividad de la muestra" se corrige a través de los pesos que se introducen en el estimador. Sin embargo, con el enfoque probabilístico, eligiendo adecuadamente el diseño existe la posibilidad de obtener muestras aceptablemente representativas, aunque ello se justifique bajo el objetivo de obtener estimadores más eficientes. En este sentido, la estratificación no sólo busca obtener estimadores más eficientes sino que además busca que la muestra sea más "representativa" de la población, o la postestratificación en la que se ajustan los pesos para que la estimación de los totales poblacionales de los post-estratos $\hat{N}_p$ coincidan con los totales poblacionales N_p, lo que se puede entender como un ajuste vía pesos en el estimador que trata de "compensar" que la muestra no sea representativa.

El muestreo probabilístico y la inferencia con base en el diseño supone que la selección de la muestra tiene lugar bajo condiciones ideales en que todos los elementos de la población tienen una probabilidad de ser incluidos en la muestra conocida y positiva. Sin embargo al llevar a la práctica los diseños probabilísticos, se observa que, debido a las limitaciones del marco o directorio y a la falta de colaboración o rechazo de algunas unidades informantes de la muestra, es muy frecuente encontrar elementos con probabilidades de inclusión nula o desconocida. En estas circunstancias, la muestra finalmente obtenida poseerá infrarepresentación de algunos grupos y en consecuencia el método aleatorio no logrará su objetivo, con el consiguiente sesgo de los estimadores establecidos a partir del diseño original.

Una ilustración de los diseños muestrales que son empleados en el área de la elaboración del índice de precios de consumo (IPC) es la siguiente: como sabemos un IPC necesita de los pesos, gastos de los hogares en bienes y servicios, los precios de estos bienes y servicios y, finalmente, la muestra de artículos sobre los que se toman los precios. Vemos que necesitamos una muestra de hogares, otra de establecimientos (puntos de consumo) y la muestra de artículos.

En España, el Instituto Nacional de Estadística usa un diseño aleatorio en la selección de los hogares que proporcionan los pesos y las parcelas de artículos, en cambio para la muestra de establecimientos emplea un diseño intencional u opinático y, finalmente, los artículos suelen ser elegidos usando el método representativo. En Suecia, el Statistics Sweden utiliza diseños aleatorios en la selección de hogares y establecimientos, y selecciona por cuotas o cutoff sampling la selección de artículos. Por último, Estados Unidos, en general, todas las muestras son aleatorias.

Las principales limitaciones del muestreo probabilístico surgen cuando existen deficiencias en el directorio o marco, cuando se produce no respuesta y en presencia de errores de medida. Veamos con más detalle la falta de respuesta y como se intentó corregirla.

La No Respuesta

Inicialmente los problemas derivados de la no respuesta se intentan resolver desde un enfoque basado en el diseño. Así por ejemplo, se supone que cada miembro de la población responderá o no con certeza generando, antes de la encuesta, el dominio o subclase de los que responde y el de los que no responden. Bajo esta hipótesis de comportamiento a la hora de responder, se elaboran técnicas para compensar el sesgo provocado por la falta de respuesta. Las que tuvieron más éxito fuero las técnicas de reponderación que modifica los pesos de los que responden, de forma que además de representarse a si mismos y a los no incluidos en la muestra representen también a los incluidos en la muestra que no responden. Un ejemplo conocido de tales técnicas es el Método de Hansen y Hurwitz (1946). En esencia la inferencia asociada con estos métodos tiene que recurrir a hipótesis o modelos que expliquen el comportamiento de los valores de las variables de interés en la población, perdiendo así su carácter de "libre de modelos".

Sobre la década de los ochenta surge lo que se denominó el enfoque de cuasi-aleatorización introducido por Oh y Scheuren (1983). Se dota a los elementos de la población de una probabilidad de respuesta, que se denomina mecanismo de respuesta o distribución de respuesta. De nuevo la inferencia basada en el diseño pierde su carácter de no dependiente de modelos, ya que para obtener inferencias validas de la muestra se requiere un determinado comportamiento de la distribución de respuesta. En la terminología de Little y Rubin (1987) la distribución de respuesta ha de ser MCAR (Missing Completely at Random), caso del estimador media ponderado, o MAR (Missing at Random), caso del estimador ponderado por clases o estimador post-estratificado. Un ejemplo lo encontramos en Thomsen y Siring (1983).

También en los ochenta, la idea de cuasi-aleatorización condujo a un enfoque en dos fases para ponderar por falta de respuesta. Ahora, el conjunto de los que responden es resultado de dos selecciones aleatoria (dos fases) sobre la población, el diseño probabilístico (conocido) que gobierna la selección de la muestra (primera fase) y el pseudo-diseño de respuesta (desconocido) que gobierna la selección del conjunto de respondientes entre la muestra (segunda fase). La diferencia con el enfoque de cuasi-aleatorización puro es que las probabilidades de respuestas están condicionadas a la muestra. Los mismos comentarios realizados para el enfoque de cuasi-aleatorización son aplicables a dos fases. Ejemplos pueden verse en el libro de Särndal y otros (1992).

El estimador ponderado por clases o el estimador post-estratificado son dos ejemplos de estimadores que utilizan pesos ajustados para tener en cuenta la falta de respuesta. Ambos consisten en dividir la muestra en clases o celdas de ponderación, utilizando información auxiliar. En esencia, partiendo de considerar que la muestra inicial es representativa, se trata de introducir ajustes en los pesos de manera que la muestra de los que responden también pueda considerarse representativa, al menos con respecto a las variables auxiliares.

4ª Aproximación: Muestras Intencionadas

Los orígenes del muestreo intencionado pueden situarse en la búsqueda de muestras equilibradas por medio de relacionar la variable de interés con un conjunto de variables de control, cuyos valores son conocidos en la población o universo considerado. Estos diseños fueron estudiados por Bowley en la memoria que realizó para el ISI en 1926. El problema surgió al producirse un sesgo en los estimadores que usaban conjuntamente los promedios de los conglomerados y un muestreo aleatorio simple de conglomerados. Ante estos sesgos, Bowley recurrió al método de regresión de Youle, que permitía relacionar una variable de interés con un conjunto de variables de control conocidas en todo el universo. El método de regresión lo usó de forma descriptiva como hoy aparece en muchos manuales de introducción a la estadística, sin apelar a los modelos de superpoblación. El muestreo era balanceado (equilibrado), donde los promedios de las variables de control en la muestra eran iguales a los promedios correspondientes de la población.

Esta idea de Bowley, de proponer un modelo de regresión entre una variable de interés y un conjunto de variables de control, para establecer una estructura estadística que permita seleccionar la muestra y asignar un error a las estimaciones, renacerá en los años 1970-71 con Royall y la introducción de los Modelos de Superpoblación.

La inferencia basada en modelos de superpoblación trata los valores de la población finita como variables aleatorias y las estimaciones se derivan del modelo especificado para estas variables. La muestra se considera fija, incluso si ha sido generada por un diseño muestral probabilística, y las inferencias están basadas en muestreo repetido del modelo de superpoblación, siguiendo un enfoque predictivo de los valores no muestrales a partir de los datos observados en la muestra. La inferencia basada en modelos fue establecida por Royall(1970, 1971) y Royall y Herson (1973). Un ejemplo reciente es el libro de Valliant, Dorfman y Royall (2000).

Notemos que la validez de los resultados, que deriva la inferencia basada en un modelo, está supeditada al grado de ajuste del modelo a la estructura real de la población. Sin embargo este problema no es nuevo ni exclusivo de la investigación en poblaciones finitas. Royall y Herson (1973) asumiendo un modelo general que expresa dependencia polinómica de la variable de interés sobre las variables auxiliares, comprueba que una muestra balanceada o equilibrada proporcionará protección frente a un error en la especificación del modelo propuesto. Para estos autores una muestra es balanceada o equilibrada de orden l cuando la media muestral de x^l es igual a la correspondiente media poblacional, para $l=1,2,\ldots,l$. De esta forma Royall y Herson recuperan o rehabilitan el muestreo intencionado para obtener una muestra balanceada o equilibrada.

Cuando la estructura estocástica asociada con los datos muestrales procede de un modelo de superpoblación adecuado a la población de interés, los resultados inferenciales los podemos analizar independientemente de que el diseño sea intencionado o aleatorio. En estas condiciones, tiene sentido plantearse la siguiente cuestión ¿existe alguna forma particular de seleccionar la muestra de manera que pueda maximizarse la eficiencia de los estimadores?

En los modelos más frecuentemente utilizados, la respuesta a esta pregunta consiste en un diseño intencionado en el que deben elegirse las unidades que verifican alguna condición particular. En particular, sobre un modelo lineal sin término independiente que vincule la variable de interés Y con una auxiliar positiva X, la precisión de las estimaciones es máxima si la muestra está formada por aquellas unidades sobre las que la variable X adopta los valores más elevados. Resultados empíricos acerca de la ventaja que aportan las muestras intencionadas pueden verse en los trabajos de Colom y otros (2006, 2007).

Al margen de las justificaciones metodológicas de algunos procedimientos de selección intencionada de la muestra, en la práctica estadística es frecuente observar mecanismos de selección intencionada que responden a intereses de tipo pragmático. En ocasiones, directamente se propone requerir información a un grupo determinado de unidades expresamente seleccionadas dentro del universo global. En el ámbito de la estadística oficial, un ejemplo lo constituye la metodología seguida por el INE en la elaboración del Índice de Producción Industrial (IPI). Los datos utilizados en la construcción del citado índice proceden de la información que intencionadamente se solicita a las empresas más grandes, aquellas que poseen el mayor volumen de producción de los productos analizados.

En estos casos no es lícito admitir que la muestra es representativa, al menos con respecto a la variable auxiliar que expresa el volumen de producción. Tampoco es admisible que las unidades muestrales puedan interpretarse como el resultado de aplicar un procedimiento de selección aleatorio. La única alternativa viable para efectuar la inferencia estadística consiste en asumir un modelo de superpoblación y basar el análisis en la estructura estocástica que éste proyecta sobre los datos muestrales.

No se trata de cuestionar la validez de la metodología utilizada para la construcción del IPI, que evidentemente es apropiada, la cuestión a resolver es ¿cómo se puede medir la fiabilidad y precisión de los índices y estimaciones construidos con dicha metodología? La respuesta a esta pregunta debe buscarse en el dominio de un modelo de superpoblación que permita evaluar la eficiencia de cada índice, considerado como un estimador inducido desde el modelo. Resultados empíricos obtenidos sobre esta cuestión han sido estudiados por Murgui y otros (2007) sobre la base de una modelo que para cada empresa relaciona la producción de un año con la producción del ejercicio que le precede.

5ª Aproximación: Muestras de Generación Espontánea

Las posibilidades que ofrecen los modelos de superpoblación para resolver problemas inferenciales sobre poblaciones finitas son más amplias de lo que en su origen cabía concebir. En los párrafos anteriores se ha puesto de manifiesto que una vez asumido un modelo, puede optarse por utilizar la estrategia óptima consistente en un estimador específico y un diseño, a veces intencionado. Pero existen otros planteamientos prácticos para los que un modelo de superpoblación ofrece soporte metodológico.

En ocasiones, el investigador no tiene la opción de intervenir en el proceso de selección muestral, estando obligado a efectuar el análisis a partir de una base de datos elaborada con la información espontáneamente proporcionada por una muestra de unidades no seleccionada previamente. Sería el caso de una petición de información remitida de manera exhaustiva a todas las unidades de un universo y sólo atendida por un número limitado. Ante una situación de este tipo caben tres posibles alternativas:

a- Considerar que la muestra es representativa y asumir que los resultados obtenidos son válidos para toda la población

b- Considerar que la muestra resultante puede interpretarse como el resultado de aplicar un proceso de selección aleatorio, ajustado a un diseño probabilístico preestablecido. En tal caso, sobre la estructura estocástica inducida por dicho diseño sería posible analizar la fiabilidad de los resultados obtenidos.

c- Asumir un modelo de superpoblación y realizar la inferencia con el soporte estocástico que se genera.

En los casos en los que existen indicios de que la falta de respuesta es no significativa, desde un punto de vista práctico y con las pertinentes reservas, podría considerarse aceptable la adopción de las dos primeras alternativas. Sin embargo, ésta no es una situación de carácter general y lo ilustraremos con un ejemplo.

Hasta 1996 la Consejería de Agricultura de la Generalitat Valenciana venía elaborando cuentas económicas agregadas anuales del sector cooperativo agrario valenciano. Para ello contaba con la información que cada año le remitían todas las entidades, si bien algunas carencias mínimas se suplían mediante técnicas de imputación. Sin embargo a partir de dicha ejercicio y debido a un cambio legislativo, las entidades han dejado de remitir sus cuentas de forma exhaustiva y las que lo hacen de manera espontánea, en ningún momento ofrecen garantías de constituir una muestra representativa del universo global de cooperativas. Antes al contrario, se observa un importante sesgo hacia las cooperativas de mayor tamaño La única alternativa viable para realizar entonces la inferencia estadística es mediante la adopción de un modelo de superpoblación.

Desde una perspectiva empírica y sobre la base de datos de las cooperativas agrarias valencianas Murgui y otros (2005) y Colom y otros (2006) utilizan la metodología de un modelo lineal, en el que los valores económicos de cada ejercicio se consideran auxiliares para la inferencia del ejercicio siguiente, para efectuar un análisis retrospectivo. La aplicación sobre los datos referidos a los años en que se disponía de información censal y muestras simuladas, ha permitido constatar que los resultados obtenidos utilizando muestras que no son representativas ni aleatorias, se ajustan con una precisión aceptable a los parámetros poblacionales, incluso con tamaños muestrales no excesivamente elevados.

Conclusiones

1ª- No existe un único procedimiento estadístico para aproximarnos al conocimiento de una población finita.

2ª- Han sido las dificultades prácticas que históricamente se han ido planteando las que han generado el desarrollo de las distintas perspectivas metodológicas.

3ª- Desde un punto de vista empírico no parece adecuado establecer un orden de preferencia entre las distintas perspectivas. Son las circunstancias particulares de cada problema las que inducen a decantarse por una alternativa frente a otra.

4ª- Uno de los problemas abiertos de mayor interés, aunque no el único, es la definición de criterios que permitan comparar el resultado de aplicar distintas aproximaciones a una población finita. En determinados contextos es frecuente disponer de diferentes estimadores y/o diseños para un mismo parámetro, resultando complejo la elección de uno de ellos en particular.

Si hubiera que concretar en una frase el objetivo común denominador en las investigaciones por muestreo, resultaría algo parecido a lo siguiente: “Se pretende diseñar las muestras y resolver los problemas de estimación minimizando los costes, maximizando la fiabilidad y precisión de los resultados y, muy especialmente, asegurando la armonización y coherencia de los resultados que se obtengan con los que se generan a través de otras operaciones estadísticas”

Bibliografía

BOWLEY, A.L. (1926): “Measurement of the Precision Attained in Sampling”, Bulletin of the International Statistical Institute, Vol. 22 (1), pp. 1-62 of Special annex following p. 451.

BREWER, K.R.W. (1999): “Design-based or Prediction-based Inference? Stratified Random vs Stratified Balanced Sampling”, International Statistical Review 67 (1), pp. 35–47.

CAMPOS F. J. Y FERNÁNDEZ DE SEVILLA (2000). *Las Relaciones Topográficas de Felipe II: Índices, fuentes y bibliografía*. Web del Real Centro Universitario. Escorial-María Cristina. San Lorenzo del Escorial.

CASSEL, C.M., SÄRNDAL, C.E. Y WRETMAN, J.H. (1976): *Foundations of Inference in Survey Sampling*, John Wiley and Sons, New York.

COCHRAN, W. G. (1953): *Sampling Techniques*, John Wiley & Sons, New York.

COLOM, C. y otros (2006): “Análisis de la evolución de una población finita con muestras repetidas”. IV Congreso de Metodología de Encuestas. (2006) Pamplona.

DEMING, W. E. (1950): *Some theory of Sampling*, Wiley, New York.

GARRIGÓS MONERRIS, J.I. (2003): “Frédéric Le Play y su círculo de reforma social”. Papers, Revista de Sociología, núm. 69, 133-146.

HANSEN, M. H., HURWITZ, W.N. Y MADOW, W.G. (1953): *Sample Survey Methods and Theory. Volume I: Methods and Applications. Volume II: Theory*, Wiley, New York.

KIAER, A. N. (1895-6): “Observations et expériences concernant des dénombrement“. Discussion appears in Liv. 1, XCIII-XCVII. The meeting was in Berne in 1895. Bulletin of the International Statistical Institute, 9, Liv. 2, 176-183.

KIAER, A. N. (1899): “Sur les méthodes représentatives ou typologiques appliquées à la statistique“. Bulletin of the International Statistical Institute, 11, Liv. 2, 180-185. The meeting was in St. Petersburg in 1897.

KIAER, A. N. (1903): “Sur les méthodes représentatives ou typologiques“. Bulletin of the International Statistical Institute, 13, Liv. 1, 66-70. Discussion The meeting was in Budapest in 1901.

KIAER, A. N. (1905): “Untitle speech on the representative method“. Bulletin of the International Statistical Institute, 14, Liv. 1, 119-126. The meeting was in Berlin in 1903.

KRUSKAL, W.H., AND MOSTELLER, F. (1980): “Representative sampling IV: The history of the concept in Statistics, 1895-1939”, International Statistical Review, 48, pp. 169-195.

LAZARSFELD, P.F. (1961): Notes on the History of Quantification in Socilogy-Trends, Sources and Problems. Isi, vol. 52, No. 2, 277-333.

LIE, E. (2002): “The rise and fall of sampling surveys in Norway, 1875-1906”. Science in Context. 15 (3), 385-409.

LITTLE, R.J.A. Y D.B. RUBIN (1987): *Statistical Analysis with Missing Data*, Wiley, New York.

Sirken, M.G. (2004): “Network sample surveys of rare and elusive populations: a historical review”. Proceedings of Statistics Canada Symposium. Innovative Methods for Surveyeing Difficult-to-reach Populations.

MURGUI, S. y otros (2005): “Diseño y evaluación empírica de una estrategia de predicción por muestreo en cooperativas agrarias”. (2005). Rev. Estadística Española. Vol 47, Número 159, 299-320.

MURGUI, S. y otros (2007): “Cálculo de la precisión en el índice de producción industrial” Congreso de Asepelt. Valladolid.

NEYMAN, J. (1934). “On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection”, Journal of the Royal Statistical Society, 97, pp. 558-606.

OH, H. Y SCHEUREN, F.J. (1983): “Weighting Adjustment for Unit Nonresponse” en W.G. Madow, I. Olkim, y D.B. Rubin (editores): *Incomplete Data in Sample Surveys: Theory and Bibliographies (Volume 2)*, Academic Press, New York, pp. 143-184.

RAO, J.N.K. (1996). “Developments in sample survey theory: An appraisal”. The Canadian Journal of Statistics, 25, 1-21.

RAO, J.N.K., AND BELLHOUSE, D.R. (1990). “History and development of the theoretical foundations of survey based estimation and analysis”, Survey Methodology, 16, pp. 3-29.

ROYALL, R.M. (1970). “On finite population sampling theory under certain linear regression models”, Biometrika, 57, pp. 377-387.

ROYALL, R. M. (1971): “Linear Regression Models in Finite Population Sampling Theory”, en V.P. Godambe y D.A. Sprott (editores): *Foundations of Statistical Inference*, Holt, Reinhart & Winston de Canada, Toronto.

ROYALL, R.M., AND HERSON, J.H. (1973): “Robust estimation in finite populations, I and II”, Journal of the American Statistical Association, 68, pp. 880-889 y 890-893.

SÄRNDAL, C.E., B. SWENSSON Y J. WRETMAN (1992): *Model Assisted Survey Sampling*, Springer-Verlag, New York.

THOMSEN, I. Y SIRING, E. (1983): “On the Causes and Effects of Nonresponse: Norweigan Experiences”, en W.G. Madow y I. Olkim (editores): *Incomplete Data in Sample Surveys: Proceedings of the Symposium (Volume 3)*, Academic Press, New York, pp. 25-29.

VALLIANT, R., DORFMAN, A.H. AND ROYALL, R.M. (2000): *Finite Population Sampling and Inference: A Prediction Approach*, New York: John Wiley & Sons, New York.

YATES F. (1946): “A review of recent developments in sampling and sample surveys (with discussion)”, Journal of the Royal Statistical Society, Vol. 109, No. 1, pp. 12-43.

YATES, F. (1949): Sampling Techniques for Censuses and Surveys, Griffin, London.

Capítulo 23

La disputa entre José Antonio de Alzate y el virrey de la Nueva España, conde de Revillagigedo, por el censo de la ciudad de México de 1790

LETICIA MAYER CELIS
Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas
Universidad Nacional Autónoma de México

Introducción

El desarrollo de los censos en la edad moderna, particularmente en el siglo XVIII, respondió a varios motivos todos ellos muy complejos: el proceso de centralización del poder, la reorganización de las monarquías, ciertos avances en medicina e higiene que propiciaron el crecimiento de la población, además de la fascinación por recopilar estadísticas con diversos tipos de información. Dentro de esta tendencia es necesario destacar el particular desarrollo que tuvieron los censos hispánicos, tanto los de la península como los de América.

Según Burke[1], en Suecia el gobierno se preocupó por estimular el crecimiento de la población, por lo que el clero estuvo obligado legalmente desde 1736 a dar cifras anuales de nacimientos y muertes en cada una de las parroquias[2] y en 1748 las autoridades ordenaron la

---

[1] Burke, Peter, 2002, pp. 176 y 177.

[2] Aunque, como bien sabemos, el interés por los nacimientos y la mortalidad se inició en 1662 cuando se publicó en Londres un libro que fue el parteaguas con el que se iniciaron los estudios estadísticos: Natural and Political Observations Mentioned in a following index, and made upon the Bills of Mortality By John Graunt. Graunt fue un comerciante de telas en Londres, cuya muy particular afición fue recolectar los boletines parroquiales en los que se hacía referencia a las personas muertas y bautizadas cada semana. Aparentemente el primer boletín se editó en 1532, año de la peste en Inglaterra. Sin embargo, la publicación continua de los mismos comenzó el 29 de diciembre de 1603. Con los datos que Graunt pudo recolectar hizo una serie de inferencias estadísticas que rompieron con viejas creencias sobre la población. Afirmó que Londres tenía 380.000 habitantes y no 6 millones como se decía popularmente, demostró que la cantidad de hombres y mujeres era proporcional, calculó el término medio de hijos por familia y desarrolló la primera tabla de vida que se conoce.

realización de un censo nacional. Con los datos que nos proporciona este autor, además de los que tengo podemos hacer el siguiente cuadro de censos en el siglo XVIII:

1748 el censo de población en Suecia

1768 primer censo español en la Península

1769 censo de Dinamarca

1777 varios censos en América como el de la Nueva España o el de Popayán

1787 segundo censo español de la Península

1790 censo de los EEUU

1790 nuevamente varios censos en las colonias hispánicas de ultramar

1801 censo del Reino Unido

1806 censo de Francia

Con lo anterior queda claro que el interés por la población, su crecimiento, distribución entre los sexos, ocupación, entre otros temas fue una preocupación del Imperio Español claramente detectada por lo menos desde el siglo XVIII. Aunque bien sabemos que Felipe II, "el rey papelero" por la cantidad de documentos que juntaba, se preocupó de los problemas de población desde el siglo XVI.

Dentro de esta problemática de los censos, la finalidad de este trabajo es exponer la discusión que se entabló entre el virrey de la Nueva España, conde de Revillagigedo, y el sacerdote criollo José Antonio de Alzate y Ramírez en torno a la veracidad de los datos del censo de la ciudad de México de 1790.[3] Dos aspectos cobran particular relevancia: 1) la inferencia que pudo hacerse con datos numéricos y 2) la discusión (activa incluso hoy en día) sobre si es necesario contar una población "cabeza por cabeza" o si es preferible calcularla.[4]

El caso de la Nueva España

En este territorio una de las discusiones sobre la población fue la que entabló Alzate con el virrey entre 1791 y 1792. Sin embargo, las reflexiones de Alzate sobre el cálculo de la población son anteriores. En 1788, en el primer volumen de la Gaceta de Literatura de México, publicó un artículo titulado: "Cálculo sobre la población de Mégico". En él afirmó que a partir de 1777 comenzaron a editarse las listas de sacerdotes que fallecieron en la ciudad de México, así como en las Guías de Forasteros las listas generales de mortalidad. Comparando ambos inventarios llegó a inferir que el número de sacerdotes muertos correspondía a la centésima parte de muertes totales. Para demostrar su afirmación hizo un pequeño cuadro comparativo de los años más significativos:

---

[3] Es importante señalar que ambos personajes debieron tener personalidades muy peculiares, por lo que la lectura de sus epístolas resulta un trabajo, no solo enriquecedor por el conocimiento científico que queda manifiesto, sino sumamente divertido por la polémica.

[4] Los documentos que hemos tomado para este análisis son parte de la correspondencia, que duró más de dos años, entre el virrey conde de Revillagigedo y Don Antonio de Alzate y Ramírez. Los textos aquí trabajados corresponden a las primeras cartas que abarcan los meses de marzo a mayo de 1791.

Años	Sacerdotes muertos	Lista general
1777	55	4585
1778	45	4446
1781	31	3894
1784	103	10463
1785	57	4971
1786	103	9112

Según Alzate, los datos de mortalidad de los clérigos eran totalmente confiables y, si la correspondencia entre la mortalidad de sacerdotes y el total en la ciudad de México era de 1 a 100, parecía fácil sacar resultados con alguna aproximación. No obstante, afirmó que este tipo de cálculos ofrecía muchos problemas. Sin embargo, su interés por conocer el número de habitantes, lo llevó a hacer afirmaciones más temerarias y decidió comparar los datos de la ciudad de México con los de Madrid. Hizo una serie de cálculos entre nacimientos y muertes, que lo llevaron a afirmar que, si la ciudad de Madrid contaba con 156.672 habitantes, la de México debería aproximarse a 210.215 vecinos.

En enero de 1791 se publicó un pequeño cuadro: "Estado General de la Población de México Capital de esta Nueva España", los datos correspondían al censo de la ciudad. En ellos se asignó la cantidad de 115.317 almas como total de los habitantes de la capital. Cuando Alzate vio estos resultados montó en cólera y escribió la primera de una serie de cartas al virrey Revillagigedo. La epístola comienza con una frase que permite imaginar la indignación del cura:

> *Trémulo tomo la pluma para informar a vuestra excelencia lo que siento [...] el estado hecho a cerca de la población de México, impreso en estos días, es de las producciones más infundadas, que han salido de la prensa.*

Los cálculos de Alzate para rebatir el censo se basaron en tres tipos de inferencias con datos diversos:

1) Comparó los planos de Madrid y México y concluyó que la segunda ciudad ocupaba una extensión en varas casi del doble que la primera, por lo que la población debería ser más abundante.

2) Siguiendo a "los más célebres calculadores políticos de la Europa, como son el Conde de Buffon, el Barón de Bielfeld", afirmó que era imposible contar una ciudad populosa, por lo que era más verosímil calcularla, lo que, según los autores antes mencionados, "se reduce a multiplicar el número de nacidos por 35, y el número de muertos por 31". Realizó estas operaciones de acuerdo a los datos con que contaba y obtuvo los siguientes resultados: la ciudad de Madrid tenía 158.480 vecinos y la de México 209.930 almas.

3) Su tercera demostración se basó en los consumos alimenticios de ambas ciudades. Sus datos le dieron casi el doble de consumo en la ciudad de México que en Madrid. A lo que el virrey contestó que los habitantes de estos reinos eran mucho más golosos que los de España.

Para Alzate estas inferencias con tres tipos de datos diferentes eran la prueba irrefutable de que la ciudad de México contaba con más de 200 mil habitantes y que, por consiguiente, estaba más poblada que Madrid, Barcelona o Sevilla. La discusión se prolongó por casi dos años. El análisis de la correspondencia nos permite ver, por una parte la disputa patriótica que encerraron los datos estadísticos y, lo que resulta más interesante, la forma de calcular la población de acuerdo a datos indirectos.

La extensión de las ciudades como una forma de inferir el número de habitantes

Alzate se sacó la espinita que el virrey le clavó al decirle que por no haber salido de la Nueva España no conocía la extensión de otras ciudades y su grandeza. "Me insinúa V.E. que vivo muy engañado en suponer a México mayor en extensión que Madrid. Yo ciertamente jamás he salido de la América" sin embargo afirmó que tenía en su poder el plano de Madrid que hizo "el erudito Don Tomás de López, y presentado al soberano por mano del excelentísimo señor Conde de Florida Blanca". Así mismo apuntó tener varios planos de la ciudad de México que hicieron otros tantos eruditos de estos reinos, "y cotejando uno con otro con el compás en la mano para medir exactamente las escalas respectivas, he notado que México ocupa una extensión de terreno considerablemente mayor que Madrid". Dentro de las mediciones, aparentemente muy exactas, que hizo Alzate sus resultados lo llevaron a afirmar que la ciudad de Madrid, del sudeste al noroeste que era la parte más larga de aquella urbe, contaba con 2500 varas menos que la de México en una medición de norte a sur. Además, consideró que en estos tramos la población estaba muy concentrada. En sus consideraciones de extensión, Alzate afirmó:

> *En Madrid por una línea que la atraviesa casi de norte a sur desde la puerta de San Bernardino al Paseo de las Delicias, el mapa representa 3600 varas. La distancia de la Puerta de Segovia a la Puerta de Alcalá es de 3000 varas. De la Puerta de Alcalá al Jardín Botánico hay 2500, de los Pozos de la Nieve a la Puerta de Embajadores Nueva, hay 3300 varas. Veamos ahora lo que México se extiende. En primer lugar de oriente a poniente: esto es, desde la garita de San Cosme a la de San Lázaro, hay 5000 varas: desde nordeste a sureste 5200 varas, y para no ser prolijo del noroeste al suroeste, hay 4600 varas.*

Nuestro autor afirmó que era probable que existieran personas que pudieran suponer que el mapa de Madrid estaba muy bien hecho y no así el de México. Sin embargo, dijo que esto sólo podían pensarlo "quienes no saben cuantos eruditos participaron y el particular empeño y cuidado con el que se trabajó en mapa de esta ciudad, con la única intención de averiguar la verdad". Para argumentar sobre la veracidad de los planos de Madrid y de México apeló a los juicios de autoridad apuntando la importancia de las personas que intervinieron en este trabajo. También recurrió a su propia experiencia, dado que Alzate levantó un plano para la distribución de los territorios parroquiales.

Nuestro autor indicó que, al argumentar sobre la mayor extensión de territorio y de población de la ciudad de México, no quiso decir que por esta razón ésta fuera más importante, "sé muy bien que ni la extensión de terreno, ni la excesiva población de una corte, deciden de su mérito". Siguió su argumentación indicando que Moscú o bien la corte del Gran Mongol son mucho más grande en población y extensión que las ciudades europeas y, sin embargo, no pueden compararse con éstas. Madrid es la capital del reino y "corte de nuestro soberano", sea cual fuere el número de sus habitantes; no obstante este hecho no justificaba que se ocultaran los hechos, ni que "se haga traición a la verdad".

Los cálculos universales basados en las listas de nacimientos y mortalidad

Para Alzate uno de los problemas fue que los habitantes del reino tenían miedo de contestar el censo. Aparentemente, el padrón que desarrolló Juan de Villalba en 1765 causó recelo entre los habitantes de la ciudad. Cuando se comenzó el nuevo censo, se pensó que los resultados serían el reclutamiento de personas para poblar las Californias, o bien someter a los hombrea a la milicia. Sin embargo, el temor a los censos no fue privativo de la Nueva España, nuestro autor anotó que en todos los países es "sumamente difícil formar un estado o padrón exacto de sus habitantes". Como ejemplo tomó lo expresado por el Barón de Bielfeld para quien un conteo exacto de población es imposible debido a la movilidad de la misma y la forma constante en que nacen y mueren los habitantes de una ciudad; siguiendo al mismo autor indicó:

> *Pero la política no necesita en este asunto de una exactitud semejante. Puede contenerse muy bien con una teórica verisimil (sic) sobre todos estos objetos, con tal que esta teórica se arrime a la verdad cuanto sea posible; y a este fin se dirigen todos los expuestos de los calculadores políticos, que debiera sostenerlos el mismo gobierno en los países bien civilizados.*
>
> *El hombre de estado se contenta con saber a poca diferencia por los extractos mortuorios, y las enumeraciones que más puedan arrimarse a la verdad, cuál es el número verisimil (sic) de habitantes que contiene el estado o gobierno a quien sirve.*
> *Es preciso repetirlo en este lugar; las listas de los niños bautizados, las de los matrimonios, y muertos forman la base de toda esta aritmética.*[5]

Es obvio que ya podemos encontrar la influencia de Jonh Graunt y sus observaciones sobre la mortalidad en Londres. Alzate citó a Buffón y apuntó la importancia de "los calculadores ingleses, franceses, holandeses, italianos que han trabajado en este asunto".[6]

Sin embargo, los datos de mortalidad de Madrid no fueron fáciles de obtener, todos los documentos con los que Alzate contó fueron las guías de forasteros de aquella ciudad, en las que

> *se da una noticia puntual y completa de los nacidos en aquella Capital, y habiendo hecho la suma de ocho años consecutivos para ver los que corresponden a cada año (porque en un año solo podía muy bien variar el número por algunas circunstancias extrañas) he conocido que en cada año nacen 4.520, que multiplicados por 35 dan el producto de 158.480 igual con corta diferencia al de 156.672, que expresa el padrón. Habiendo sumado igualmente el número de nacidos en México en otros ocho años veo le corresponden en cada año 5 998, que multiplicados por 35 dan el producto de 209.930. De donde se infiere que en México hay 51.450 habitantes más que en Madrid.*

[5] La cita que hizo Alzate de Bielfeld, que es mucho más larga que la que hemos apuntado, la sacó, según él apuntó, de las Instituciones Políticas, pp. 196-200. En las cartas al virrey, Alzate no nos da la cita exacta, pero sí indicó que esta obra se tradujo al castellano y se dedicó al Sr. Conde de Aranda. Esta afirmación provocó la ira del virrey que se dejó ver en las siguientes misivas.

[6] La lectura cuidadosa de esta correspondencia no deja de sorprenderme, al darme cuenta de que Alzate estaba muy bien enterado de los escritos y discusiones europeas sobre los cálculos poblacionales. Es muy probable que sus fuentes directas no fueran Graunt, Pascal, Leibniz, Fermat o Huygens, pero creo que conoció los trabajos de todos estos autores a través de algunos otros más modernos (del siglo XVIII y no del XVII) que retomaron los planteamientos y los profundizaron.

Aunque, como se verá más adelante, Alzate completó sus cálculos, no solo con los datos de natalidad, sino con las listas de mortalidad de la ciudad de México, es claro que sus inferencias no pueden ser muy válidas. Sin embargo no lo fueron, en general, las que se hicieron en aquella época. No obstante, la validez de la argumentación está en plantear el problema de calcular a la población a través de datos indirectos, en vez de contarla cabeza por cabeza.

Los cálculos de mortalidad se complicaron mucho más. Por una parte no contó con las listas de mortalidad de Madrid, por otro lado, aunque sí tuvo las listas de mortalidad de la ciudad de México, se enfrentó a la crítica de sus contemporáneos. Estos afirmaban que el número de muertes en el Nuevo Mundo era mucho mayor que en Europa, por lo que no era válido seguir los universales planteados por Buffón de multiplicar los muertos por 31. Para poder demostrar que la relación era igual en México que en Europa recurrió a los casos de mortalidad en colegios de la ciudad de México, en la Parroquia de Santo Tomás que fue de indios que habitaban el sur de la ciudad y de la Real Fábrica de Cigarros. Cada caso lo tomó por separado y sus cálculos apuntaron a que de 30 individuos, sólo 1 moría. En otras palabras la proporción de Buffón se cumplía con la misma exactitud en el viejo mundo, como en el nuevo. En su reflexión, para argumentar en contra del censo, afirmó lo siguiente:

> Ahora bien si entre estos indios que llevo referidos, que seguramente es la gente más infeliz, así por su modo de vivir, como por su habitación, y otras varias circunstancias, que V.E. no ignora, es tan corto el número de los muertos: si los que se hayan empleados en la Fábrica de Tabaco, y se ven precisados a respirar un aire cargado de exhalaciones tan perniciosas a la salud, como las que exhala el tabaco, mueren en una proporción tan corta, como la insinuada ¿qué diremos de los que viven en el centro de la ciudad, se alimentan de viandas más inocentes, y se hallan ocupados en agencias incomparablemente menos peligrosas?

Sin embargo Alzate no se quedó con esta sola reflexión. Apuntó la importancia de la mortalidad infantil, y el que la aritmética política del mundo afirmaba que muchos de los nacidos no llegaban al año de vida. En el caso de la Nueva España, algunos autores podían suponer que la mortalidad de niños era mayor que la de Europa.

> *¿Qué importa, dicen algunos, que sea mayor el número de los nacidos, si la mayor parte de estos perece regularmente en el primer año de su nacimiento? Mas para desvanecer esta cavilación basta decir, que he registrado los libros relativos a los expósitos, que entraron y murieron en estos dos últimos años, y vi con admiración que habiendo entrado de todas castas 212, solamente murieron 79, esto es, casi la tercia (sic) parte de los que entraron, cuando en Europa mueren ordinariamente la mitad.*
>
> *De aquí deduzco una reflexión que naturalmente se presenta aún al menos advertido. Se sabe que a la cuna solo se conducen aquellos niños que por lo regular han sido el fruto de la disolución de sus madres. Éstas para ocultar su preñez naturalmente procuran ceñirse mucho y aún toman varias bebidas para abortar, y cuando no consiguen de este modo su intento, el feto, que aún se halla demasiado tierno y débil, no puede menos que sentir una novedad, y alteración considerable. Llega el tiempo del parto y si esta infeliz criatura fue concebida en temor, nace igualmente entre el susto la turbación y el espanto (...)*[7]

[7] Algunos meses después, en septiembre de 1791, Alzate publicó en la Gaceta de Literatura un artículo sobre el cálculo de la población de todo el mundo. Es una traducción del portugués; en él se habla de las tendencias

Para Alzate, la argumentación de que incluso los niños expósitos tienen más posibilidades de sobrevivir en México que en Europa fue una prueba irrefutable de que, por lo menos, podía calcularse el crecimiento de la población con los mismos índices que en Europa y utilizando los escritos de Buffón.

Dentro de esta argumentación retomó la mortalidad entre los religiosos y religiosas. La importancia de este rubro radicó en que le sirvió como "grupo de control", en términos actuales. Los religiosos de la ciudad de México resultaban fáciles de contar. Se sabía cuántos había en cada convento, parroquia o iglesia. También se sabía cuando morían o si eran transferidos a algún otro lugar. Además, según afirma nuestro autor, se publicaban listas anuales de los sacerdotes o religiosos y religiosas que morían. Como apuntamos anteriormente, Alzate calculó que las defunciones de los religiosos de la ciudad de México, comparada con el resto de la población, fue de 1 a 100; lo que según él le permitió calcular la mortalidad general.

Los consumos alimenticios como una manera de calcular la población

En este sentido, lo primero que hizo Alzate fue una lista de varias especies de alimentos específicos de México. En ella apuntó ocho tipos diferentes de carne, veintinueve variedades de vegetales, veinticinco tipos de frutas y dos bebidas. Una vez hecha la lista, averiguó la cantidad que se vendía de cada uno de estos productos. También hizo la aclaración de que había muchos alimentos que se vendían "ocultamente", por lo que consideró que sus cálculos eran en realidad más bien conservadores.

> *Por lo que se acaba de decir se infiere con evidencia que el consumo de México aun limitándose únicamente a las partidas insinuadas no baja de 109,085,000 libras. Ahora si se quiere saber lo que corresponde a cada individuo diariamente, suponiendo la población de 200 000 habitantes, fórmese el cálculo y se verá que no baja de una libra y siete onzas sin excepción de niños, enfermos, que es una cantidad excesiva (...) Pero no es esto todo. Nos falta aun que calcular el chocolate, y las semillas, que se gastan diariamente (...) las gallinas, huevos, queso, leche (...) no dudo afirmar o que es preciso que en México cada individuo consuma diariamente casi tres libras sin excepción de niños y enfermos, o es menester convenir que la población de México asciende a más de 200 000 individuos.*

Alzate afirmó que ante la contundencia de sus argumentos ha habido algunos individuos que suponen "que en el reino son menos substanciosos los alimentos", por lo que se debe de comer más que en Madrid. Sin embargo afirma que no existen bases para "proferir una proposición tan extraña".

Nuestro autor comparó sus listas de alimentos con algunas que le llegaron de Madrid. En este ejercicio resultó que el número de carneros que se introducían en aquella "villa para su abasto apenas llegaba a 200.000...cuando en México ha habido año en que se hayan (sic) matado 450.000...esto es 250.000 carneros más que en Madrid". Sin embargo, Alzate reconoce que no tiene los suficientes datos como para poder comparar, con más exactitud, los consumos de ambas ciudades.

---

universales en el crecimiento de la población, la alta mortalidad infantil y la esperanza matemática de la vida. Es de suponer que Alzate lo publicó como información y sustento a su refutación al censo.

Algunas consideraciones

La correspondencia entre Alzate y el Conde de Revillagigedo sobre este asunto fue mucho más larga y complicada que lo apuntado en esta presentación. El problema se llevó dos años de epístolas y varias publicaciones de diferentes tipos. ¿Quién tuvo la razón?, nunca lo vamos a saber. Lo importante es apuntar que en la Nueva España existieron “sabios” que, sin lugar a dudas, estaban informados de los progresos y publicaciones sobre el cálculo estadístico, la esperanza matemática y la ya casi bizantina discusión sobre qué es más exacto cuando se trata de conocer a una población: contar o calcular.

Capítulo 24

Una explicación de las regularidades detectadas por Pascal en su tabla de valores de las partidas

Francisco Javier Ortega Irizo
José Antonio Camúñez Ruíz
Jesús Basulto Santos
Universidad de Sevilla

Introducción

Entre el verano y el otoño de 1654 tuvo lugar la correspondencia entre los sabios franceses Blaise Pascal (1623-1662) y Pierre de Fermat (1601-1665), donde se abordan, además de otros problemas sobre aritmética y geometría, la resolución de tres problemas sobre juegos de azar: El problema del dado con partidas no jugadas, el problema de los dados del Caballero de Méré y el Problema de los Puntos, siendo este último el que consideraremos en nuestro trabajo.

De forma resumida, el Problema de los Puntos se presenta en la siguiente situación: dos jugadores, A y B, se ponen de acuerdo en jugar una serie de juegos justos, que llamaremos partidas, hasta que uno de ellos haya ganado un número especificado de ellas, por ejemplo, n. Por alguna razón, los jugadores deciden interrumpir el juego cuando el jugador A ha ganado g_a partidas y B ha ganado g_b, donde g_a y g_b son menores que n, o sea, el juego ha sido interrumpido antes de la conclusión del mismo. Si los jugadores A y B han apostado la misma cantidad monetaria, k, surge entonces el problema de cómo repartir la cantidad total apostada, $2k$, entre ambos. La forma de hacer el reparto es lo que se conoce como Problema de los Puntos.

Aunque hubo soluciones anteriores a este problema (en algunos manuscritos medievales de Italia y Francia y las proporcionadas por los matemáticos italianos del Renacimiento como Pacioli, Tartaglia, Forestani, Peverone o Cardano) las más conocidas son las que aportaron Pascal y Fermat en su famosa correspondencia de 1654 y en la publicación póstuma, en 1665, del Triángulo Aritmético de Pascal. Estos autores lo abordan acotando el mismo dentro de los siguientes supuestos: (1) Por juego justo entienden que los jugadores tengan la misma habilidad y, también, que sus apuestas sean proporcionales a sus habilidades, es decir, deben apostar la misma cantidad monetaria. En lenguaje actual, el supuesto sería que la probabilidad de que el jugador A gane una partida es igual a 0.5. (2) Los jugadores "*ni se cansan ni aprenden en el desarrollo del juego*", es decir, sus habilidades permanecen constantes y, así, no les afectan los resultados obtenidos en partidas anteriores. Este último supuesto corresponden a lo que entendemos en el campo de las probabilidades como independencia, es decir, supone que los distintos juegos son conjuntamente independientes. Por último, el supuesto (3) se refiere a que al interrumpirse el juego la información relevante es el número de partidas que le restan por ganar a cada jugador y, así, están suponiendo que el juego interrumpido puede reanudarse hasta finalizar con un ganador. La situación existente al interrumpirse el juego la representaremos por (a,b), donde los valores a y b son el número de partidas que le restarían por ganar a los jugadores A y B respectivamente para ser proclamados vencedores del juego.

Bajo los tres supuestos anteriores, Pascal y Fermat resuelven el problema de los puntos mediante los siguientes métodos: (a) Fermat sumerge el problema de los puntos en un nuevo problema con un número máximo de partidas y lo resuelve mediante combinatoria; (b) Pascal propone su regla de la esperanza y a partir de su trabajo sobre el Triángulo Aritmético, que hoy denominamos, Triángulo de Pascal, resuelve el problema a partir de la binomial simétrica y (c), ante las dudas que planteó Pascal a su primera propuesta, Fermat resuelve el problema a partir de lo que hoy denominamos binomial negativa.

Para el caso genérico de juego interrumpido en la situación (a,b), es decir, al jugador A le faltan a partidas y b al jugador B, si $2k$ es el total apostado por ambos jugadores, Pascal calcula la probabilidad de que el jugador A ganase el juego, si éste continuase, por la siguiente fórmula:

$$P_A\left[(a,b)\right]=\sum_{i=a}^{a+b-1}\binom{a+b-1}{i}\left(\frac{1}{2}\right)^{a+b-1},$$

y, por tanto, dicho jugador debe recibir la cantidad $2k\cdot P_A\left[(a,b)\right]$.

Uno de los problemas particulares que aborda Pascal dentro de este contexto es el siguiente. Partiendo de la situación (b,b), es decir, a ambos jugadores les restan el mismo número de partidas para ganar el juego, consideramos los juegos $\left\{(b-j,b);\quad j=1,2,...,b\right\}$, o sea, juegos en los que el jugador A va ganando partidas de forma sucesiva, mientras que B no gana ninguna. Bajo esta premisa, Pascal compara las situaciones $(b-r+1,b)$ con $(b-r,b)$, es decir, el jugador A al que le faltaban $b-r+1$ partidas ha ganado una más (la partida *r*-ésima) y, por tanto, ahora le faltan $b-r$ partidas. El interés de Pascal consiste en calcular el valor que tiene la partida *r*-ésima ganada por el jugador A, valoración que se hace sobre la apuesta del otro jugador. O sea, Pascal calcula la proporción de apuesta del jugador B que le aporta al jugador A al ganar la partida *r*-ésima. Si llamamos a esta proporción $W_A\left[(r,b)\right]$, entonces por ejemplo, $W_A\left[(2,b)\right]$ es la diferencia entre la proporción que el primer jugador se lleva del otro al haber alcanzado el juego $(b-2,b)$ y esta misma cuando llegó al juego $(b-1,b)$. En este caso, usando el lenguaje de Pascal, diríamos que $W_A\left[(2,b)\right]$ mide "*el valor de la segunda partida*". Pascal demuestra el siguiente resultado:

$$W_A[(r,b)] = \binom{2b-r-1}{b-r}\left(\frac{1}{2}\right)^{2b-r+1},$$

cuya demostración con notaciones más actuales puede verse en Basulto y Camúñez (2007,a).

Esta fórmula le permite calcular $W_A\left[(r,b)\right]$ para distintos valores de b y r. En su carta del 29 de julio 1654, Pascal incluye la siguiente tabla de valores de las sucesivas partidas ganadas por el primer jugador como proporción de lo apostado por el segundo.

Tabla I. Tabla de Pascal:Valor de las partidas

		Número de partidas del juego					
		6	5	4	3	2	1
Valor de la partida	1ª	0’2461	0’2734	0’3125	0’375	0’5	1
	2ª	0’2461	0’2734	0’3125	0’375	0’5	
	3ª	0’2188	0’2344	0’250	0’250		
	4ª	0’1641	0’1563	0’125			
	5ª	0’0938	0’0625				
	6ª	0’0313					

En esta tabla las filas son los valores de las partidas, de la primera a la sexta, y las columnas corresponden a valores de *b*. Pascal presenta esta tabla multiplicada por 256 para operar con enteros, nosotros hemos trabajado con los valores calculados de las partidas. Sobre esta tabla, Pascal escribe a Fermat en la citada carta:

> *“Veréis igualmente que los números de la primera línea aumentan siempre, lo mismo los de la segunda, lo mismo los de la tercera. Pero a continuación, los de la cuarta disminuyen, los de la quinta, etc. Esto es lo que resulta extraño”.*

Igual que Pascal se extrañaba de las regularidades encontradas en su tabla, a nosotros nos extraña que ninguno de los autores que hemos investigado, posteriores a Pascal y Fermat, hayan abordado esta cuestión más particular que se presenta de forma colateral a la resolución del Problema de los Puntos.

Por ejemplo, Huygens (1629-1695) recoge, en su conocida obra *De Ratiociniis in Ludo Alea* (1657, versión latina, 1660, versión holandesa, la primera obra impresa sobre este cálculo), los resultados de Pascal y Fermat sobre la resolución del Problema de los Puntos, sin añadir nada nuevo sobre este asunto. Igual ocurrirá con Jacob Bernoulli que en su obra, The Art of Conjecturing, publicada en 1713, y cuya traducción en el 2006, se debe a Edith Dudley Sylla, recoge en las páginas 106-110 los mismos resultados que Huygens, a pesar que en un trabajo sobre *Lettre a un Amy, sur les Parties du Jeu de Paume*, Jacob Bernoulli introduce jugadores con distintas habilidades en un juego de Tenis.

Para el caso de que las habilidades de los jugadores sean distintas, sean estas *p* y *q*, con $p+q=1, \quad p>0, \quad q>0$, si mantenemos que las probabilidades son proporcionales a dichas habilidades en cada partida, Johann Bernoulli, un primo de Jacob Bernoulli aportará una generalización de la fórmula de Pascal para el cálculo de la probabilidad de que el jugador A gane el juego, si éste continuase:

$$P_A\left[(a,b)\right] = \sum_{i=a}^{a+b-1}\binom{a+b-1}{i}p^i q^{a+b-1-i}.$$

Esta fórmula aparece, sin demostración, por primera vez en una carta de fecha 17 de marzo de 1710 que Johann Bernoulli envió a P. R. Montmort, el cual la recoge en la segunda edición de su obra, *Essay d'analyse sur les jeux de hazard, 1713 second edition,* en las páginas 232-248. En este fragmento, Montmort prueba esta fórmula de Johann Bernoulli, añadiendo otra equivalente que generaliza la solución última de Fermat mediante la Binomial Negativa (Basulto el al, 2005). De forma independiente, A. de Moivre resolverá el problema de los puntos para desigual habilidad de los jugadores y para más de dos jugadores en sus obras de *De Mensura Sortis* (1712) y *The Doctrine of Chances*.

Como ya se ha dicho, en ninguna de las obras citadas que siguieron al trabajo de Pascal hemos encontrado el estudio y valoración individualizada de cada partida ganada por un jugador. Así pues, el objeto del presente trabajo es explicar el "*comportamiento extraño*" detectado por Pascal en base a propiedades de la distribución de probabilidad binomial (obsérvese que $W_A[(r,b)] = P\left[X_{2b-r-1} = b-r\right]$, donde X_{2b-r-1} es una variable aleatoria que sigue un modelo binomial de con $2b-r-1$ pruebas independientes y probabilidad de éxito $p = 0.5$). Asimismo, extenderemos los resultados obtenidos por Pascal para el caso de jugadores con igual habilidad a la situación genérica en que las habilidades de los jugadores sean p y q, donde ($p+q=1$).

A partir de aquí, en la sección 2 generalizamos la fórmula de valoración de las partidas al caso de jugadores con distinta habilidad; en la sección 3 analizamos las propiedades de la distribución binomial que nos permitirán, en la sección 4, explicar el comportamiento de la tabla construida por Pascal. En la sección 5, exponemos las principales conclusiones de este trabajo.

Valoración de las partidas ganadas en el caso de jugadores con distinta habilidad.

Para un juego interrumpido en la situación (a,b), donde suponemos que las habilidades de los jugadores A y B son p y q, respectivamente, que son valores del intervalo $(0,1)$, con $p+q=1$, la probabilidad de que el jugador gane el juego, suponiendo que éste continuase, como ya hemos señalado viene dada por la fórmula

$$P_A\left[(a,b)\right] = \sum_{i=a}^{a+b-1} \binom{a+b-1}{i} p^i q^{a+b-1-i}.$$

Ahora, nuestro interés se centra en obtener el valor de cada partida en la situación descrita en el epígrafe primero, cuando las habilidades de los jugadores son distintas.

En este caso, no supondremos que cada jugador apuesta una cantidad monetaria *K*, sino que trabajaremos bajo el supuesto de que cada jugador hace una apuesta proporcional a su habilidad. El hecho de trabajar bajo esta suposición de apuesta proporcional a la habilidad, nos va a permitir generalizar los resultados obtenidos por Pascal con la binomial simétrica (de parámetro $p = 0'5$) a la binomial general de parámetro p. O, en otras palabras, los resultados obtenidos por Pascal para el caso simétrico pueden generalizarse al caso general siempre que supongamos que las apuestas sean proporcionales a las habilidades. Concretamente, sin pérdida de generalidad, supondremos que el total apostado es la unidad, con lo que las cantidades apostadas por los jugadores A y B serán p y q respectivamente.

En esta situación, el siguiente lema generaliza la relación encontrada por Pascal entre los valores de las distintas partidas con las probabilidades de una variable aleatoria binomial.

Lema 1: $W_A[(r,b)] = \binom{2b-r-1}{b-r} p^{b-r} q^{b-1}$.

La demostración de este lema se recoge en el apéndice I.

Como podemos comprobar, el valor de la partida r-ésima se corresponde con la probabilidad de que una variable aleatoria binomial con $2b-r-1$ pruebas independientes cada una de ellas con probabilidad de éxito p tome el valor $b-r$, es decir, $W_A[(r,b)] = P[X_n = b-r]$, donde $X_n \sim Bi(n,p)$, siendo $n = 2b-r-1$.

Pascal encontró el comportamiento "*extraño*" al analizar su tabla por filas. Así, lo que hace es comparar el valor de una de las partidas (por ejemplo la primera, la segunda, etc.) cuando el número inicial de partidas que le restaban a cada jugador era b y $b-1$ respectivamente. Es decir, Pascal se interesa por las comparaciones del tipo $W_A[(r,b)]$ con $W_A[(r,b-1)]$. Él encuentra que cuando $r = 1,2,3$, $W_A[(r,b)]$ siempre es mayor (en un caso igual) que $W_A[(r,b-1)]$, mientras que se da la situación contraria para $r = 4,5$. Puesto que el valor de cada par (r,b) puede identificarse con la probabilidad de un valor concreto de una variable binomial con probabilidad de éxito p, utilizaremos esta correspondencia para explicar el comportamiento de la tabla de Pascal y generalizarlo también al caso de distintas habilidades. Ahora bien, el valor $b-r$, por conveniencia, vamos a expresarlo en la forma $\frac{n}{2} - l$ o $\frac{n-1}{2} - l'$, para lo que resulta conveniente distinguir los casos en que r es par de aquellos en los que es impar.

Concretamente, cuando r es impar, entonces $n = 2b-r-1$ es par y en este caso nos interesa relacionar el par (r,b) con la probabilidad del valor $\frac{n}{2} - l$ de X_n; ahora bien, como sabemos que dicho valor es $b-r$, entonces debe cumplirse $\frac{n}{2} - l = b-r$, y por tanto $l = \frac{r-1}{2}$, pudiendo escribir ahora $W_A\left[(r,b)\right] = P\left[X_n = \frac{n}{2} - l\right]$, donde $n = 2b-r-1$ y $l = \frac{r-1}{2}$. Por ejemplo, la primera partida, $r = 1$, la relacionamos con el valor $\frac{n}{2}$, que al ser n par ($n = 2b-2$) corresponde al valor mediano o central del conjunto de posibles valores que toma la variable X_n, es decir, las situaciones $\{(1,6),(1,5),(1,4),(1,3),(1,2)\}$ se corresponden siempre con los valores $\{5,4,3,2,1\}$ de una distribución binomial en la que n va tomando los valores $\{10,8,6,4,2\}$ respectivamente. Para la partida tercera, $r = 3$, $l = 1$, y, por tanto, el valor relacionado es $\frac{n}{2} - 1$, el primer valor a la izquierda del valor central ($n = 2b-4$), y así sucesivamente.

Para las partidas pares, $r = 2,4,6,\ldots$, se tendrá que $n = 2b-r-1$ es impar y ahora relacionamos la partida r-ésima con el valor $b-r = \frac{n-1}{2} - l'$ de una variable binomial con parámetros (n,p), donde $l' = \frac{r-2}{2}$ y, por el Lema I, $W_A\left[(r,b)\right] = P\left[X_n = \frac{n-1}{2} - l'\right]$. Por ejemplo, la segunda partida, $r=2$, la relacionamos con el valor $\frac{n-1}{2}$, que al ser n impar ($n = 2b-3$) corresponde al menor de los dos valores centrales del conjunto $\{0,1,2,...,n\}$, valores que toma la variable X_n. Para la partida cuarta, $r=4$, tendremos $n = 2b-3$ y el valor relacionado es $\frac{n-1}{2} - 1$, y así sucesivamente.

La anterior correspondencia entre partidas y valores de una variable binomial, no superiores al mínimo de los valores medianos, permite pasar de estos valores de una variable binomial a las partidas de juegos interrumpidos. Así, si consideramos la familia de variables binomiales (n,p), con *n* par, los valores $X_n = \frac{n}{2} - l$, con $l = 0,1,...,\frac{n}{2}$, los podemos relacionar con las partidas impares $r = 2l+1$, para juegos con $b = n+1-\left(\frac{n}{2} - l\right)$, siendo $P\left[X_n = \frac{n}{2} - l\right] = W_A\left[(r,b)\right]$; mientras que si consideramos la familia de variables binomiales (n,p), con *n* impar, los valores $X_n = \frac{n-1}{2} - l'$, con $l' = 0,1,...,\frac{n-1}{2}$, los podemos relacionar con las partidas pares $r = 2l'+2$, para juegos con $b = n+1-\left(\frac{n-1}{2} - l'\right)$, siendo $P\left[X_n = \frac{n-1}{2} - l'\right] = W_A\left[(r,b)\right]$.

Propiedades de las probabilidades de la familia binomial (n,p) con *n* par detectadas en los valores $X_n = \frac{n}{2} - l$, con $l = 0,1,...,\frac{n}{2}$.

En este epígrafe consideraremos dos situaciones del tipo (r,b) y $(r,b-1)$. Para comparar los valores de estas partidas, dados por $W_A[(r,b)]$ y $W_A[(r,b-1)]$, establecemos la correspondencia de cada partida con un valor concreto de una distribución binomial, y entonces procederemos a comparar las probabilidades de ambos valores. Vamos a considerar inicialmente valores de r impares (que se asocian a distribuciones binomiales de parámetro par), siendo todo el desarrollo similar en el caso de partidas pares.

Como hemos descrito, el valor de la partida (r,b) cuando r es impar, es igual a la probabilidad del valor $x = \frac{n}{2} - l$ (donde $l = \frac{r-1}{2}$), en una distribución binomial de parámetros $n = 2b-r-1$ y p. Así, $W_A[(r,b)] = P\left[X_n = \frac{n}{2} - l\right]$; análogamente, el valor de la partida $(r,b-1)$ será la probabilidad del valor $x' = \frac{n'}{2} - l$ en una distribución binomial de parámetros

$n' = 2b - r - 3$ y p. Así, $W_A[(r,b-1)] = P\left[X_{n'} = \frac{n'}{2} - l\right]$; Puesto que $n' = n-2$, también tendremos $x' = \frac{n-2}{2} = x - 1$. De esta forma, la comparación de $W_A[(r,b)]$ con $W_A[(r,b-1)]$, podemos llevarla a cabo comparando $P[X_n = j]$ con $P[X_{n-2} = j-1]$, donde n es un entero par. Para comparar estas dos probabilidades, consideramos el polinomio

$$P_n(j) = \frac{P\left[X_{n-2} = j-1\right]}{P\left[X_n = j\right]} - 1 = \frac{\binom{n-2}{j-1} p^{j-1} q^{n-j-1}}{\binom{n}{j} p^j q^{n-j}} - 1 = \frac{j(n-j)}{n(n-1)pq} - 1,$$

y ahora tendremos qué ver cuándo este polinomio es mayor, menor o igual a 0.

Las raíces de este polinomio son $\frac{n}{2} \pm \lambda_n$, donde $\lambda_n = +\sqrt{\frac{n^2}{4} - pqn(n-1)}$; Para el problema que nos ocupa, solo nos interesa el recorrido $0 \le j \le n/2$, ya que los valores que se corresponden con la valoración de las partidas siempre son menores o iguales que $n/2$. Es fácil comprobar que el polinomio es negativo en el intervalo $[0, n/2 - \lambda_n)$, positivo en el intervalo $(n/2 - \lambda_n, n/2]$ e igual a cero en el punto $n/2 - \lambda_n$. Diremos que un valor de la distribución binomial está "*en la zona central*" si pertenece al intervalo $(n/2 - \lambda_n, n/2]$; el valor está "*en la cola*" si pertenece al intervalo $[0, n/2 - \lambda_n)$; si el valor coincide con $n/2 - \lambda_n$, diremos que es "*frontera*".

Así, podemos encontrarnos tres situaciones, que recogemos enunciadas en forma de lema (de demostración inmediata).

Lema 2: En la situación descrita con anterioridad, se verifica:

i) Si $j < n/2 - \lambda_n$, entonces $P_n(j) < 0$, es decir, $P\left[X_{n-2} = j-1\right] < P\left[X_n = j\right]$. Como consecuencia, obtenemos que si el valor asociado a la partida (r,b) está "*en la cola*", entonces $W_A[(r,b-1)] < W_A[(r,b)]$, es decir, al disminuir b, disminuye el valor de la partida $r - \acute{e}sima$.

ii) Si $n/2 - \lambda_n < j \le n/2$, entonces $P_n(j) > 0$, es decir, $P\left[X_{n-2} = j-1\right] > P\left[X_n = j\right]$. Como consecuencia, obtenemos que si el valor asociado a la partida (r,b) está "*en la zona central*", entonces $W_A[(r,b-1)] > W_A[(r,b)]$, es decir, al disminuir b, aumenta el valor de la partida $r - \acute{e}sima$.

iii) Si $j = n/2 - \lambda_n$, entonces $P_n(j) = 0$, es decir, $P\left[X_{n-2} = j-1\right] = P\left[X_n = j\right]$. Como consecuencia, obtenemos que si el valor asociado a la partida (r,b) está "*en la frontera*", entonces $W_A[(r,b-1)] = W_A[(r,b)]$, es decir, al disminuir b, se mantiene el valor de la partida $r - \acute{e}sima$.

Ejemplo 1: En el caso de la primera partida, $r=1$, sabemos que los valores asociados siempre son de la forma $n/2$, por lo que siempre están "*en la zona central*", y por tanto, el valor de la primera partida va aumentando a medida que b disminuye (como observó Pascal en su tabla para el caso $p=0'5$).

Ejemplo 2: Vamos a considerar el caso $p=0'65$, y la partida quinta (es decir, $r=5$), cuando inicialmente restaban por ganar 5,6,7,8 y 9 partidas (es decir, consideramos los valores $b\in\{5,6,7,8,9\}$). Comparemos la valoración de la partida quinta para cada uno de los valores de b.

Tabla II: Valor de la partida quinta ($b\in\{5,6,7,8,9\}$, $p=0'65$*).*

b	9	8	7	6	5
$n=2b-r-1=2b-6$	12	10	8	6	4
$x_n=\frac{n}{2}-\frac{r-1}{2}=\frac{n}{2}-2$	4	3	2	1	0
$\frac{n}{2}$	6	5	4	3	2
$\frac{n}{2}-\lambda_n$	3'557	2'872	2'194	1'525	0'873
$P\left[X_n=x_n\right]=W_A\left[(r,b)\right]$	0'0199	0'0212	0'0217	0'0205	0'0150

Para la partida $(5,9)$, vemos que el valor $x_n=4$ está "*en la zona central*", pues pertenece al intervalo $\left(n/2-\lambda_n, n/2\right]=\left(3'557,6\right]$; estamos en la situación **ii)** y por tanto, al disminuir b, el valor de la partida quinta debe aumentar, como se observa en la tabla, ya que el valor pasa de 0'0199 a 0'0212. Para la partida $(5,8)$, vemos que el valor $x_n=3$ también está "*en la zona central*", pues pertenece al intervalo $\left(n/2-\lambda_n, n/2\right]=\left(2'872,5\right]$; de nuevo estamos en la situación **ii)** y por tanto, al disminuir b, el valor de la partida quinta debe aumentar, como se observa en la tabla, ya que el valor pasa de 0'0212 a 0'0217. Sin embrago, en la situación $(r,b)=(5,7)$, el valor $x_n=2$ queda "*en la cola*" pues no pertenece al intervalo $\left(n/2-\lambda_n, n/2\right]=\left(2'194,4\right]$, por lo que estaríamos en la situación **i)** y así el valor de la partida quinta debe disminuir al disminuir b, como también podemos observar en la tabla. Idéntica situación ocurre para la partida $(5,6)$.

Aplicación al caso de la binomial simétrica: explicación del comportamiento de la tabla de Pascal.

Vamos a aplicar los resultados obtenidos en la sección 3 al caso particular de la binomial simétrica ($p=0'5$) y a los valores de r y b analizados por Pascal en su tabla, al objeto de explicar el "*comportamiento extraño*" observado por el autor.

Indiquemos en primer lugar que para el caso $p=0'5$, el polinomio que hemos usado para comparar las probabilidades sería

$$P_n(j)=\frac{4j(n-j)}{n(n-1)}-1,$$

cuyas raíces son $n/2 \pm \lambda_n$, donde $\lambda_n = +\sqrt{n}/2$; Así, el punto "*frontera*", que distingue la "*zona central*" de la cola sería en este caso $n/2 - \sqrt{n}/2$ (que coincide con la esperanza de la distribución menos una vez la desviación típica de la misma). Comencemos con el caso de las partidas impares (es decir, $r = 1,3,5$). El comportamiento de la partida primera, ya ha sido explicado en el ejemplo 1 de la sección tercera. Recordemos que en este caso los valores asociados coinciden siempre con $n/2$ y por tanto siempre están en la "*zona central*" (situación **ii)** del Lema 2), con lo que el valor de la partida primera aumenta a medida que disminuye b.

Para la tercera partida, cuando $b = 6$, tenemos que considerar el valor $x = 3$ en la distribución binomial para $n = 8$. En esta distribución, el punto "*frontera*" es 2'586, por lo que estaríamos en la "*zona central*" (situación **ii)**) y por tanto la probabilidad debe aumentar al pasar al caso $b = 5$, como efectivamente ocurre en la tabla. Idéntica situación se da al analizar el juego $(3,5)$, en el que $n = 6$, $x = 2$ y el punto "*frontera*" es 1'7753. Sin embargo, en el juego $(3,4)$, en el que $n = 4$, el valor de la binomial correspondiente, $x = 1$, coincide con el punto "*frontera*" , por lo que estaríamos en la situación **iii)** del Lema 2 y por tanto la probabilidad (o valoración de la partida) debe mantenerse al pasar al valor $b = 3$; efectivamente, vemos que en los casos $b = 4$ y $b = 3$ la valoración de la tercera partida es igual a 0'25.

Para la quinta partida, sólo tiene sentido ocuparse del caso $b = 6$. Este juego se corresponde con el valor $x = 1$ de una distribución binomial simétrica con $n = 6$, en la que el "*punto frontera*" sería 1'7752; así, el valor asociado está "*en la cola*" de la distribución (situación **i)** del Lema 2), por lo que el valor de la partida quinta debe disminuir al pasar al caso $b = 5$, como efectivamente ocurre.

Aunque solo hemos analizado el caso de partidas impares (que se corresponden con binomiales en las que n es par), todos los resultados pueden extenderse sin mayor dificultad cuando la partida sea par, teniendo en cuenta las observaciones efectuadas en la sección segunda. Vamos a considerar el caso concreto de la tabla de Pascal, para explicar también el comportamiento de las partidas pares ($r = 2,4$). En este caso, el intervalo que define la "*zona central*" sería $\left(\frac{n}{2} - \sqrt{\frac{n^2}{4} - pqn(n-1)}, \frac{n-1}{2}\right]$ (observemos que las raíces del polinomio $P_n(j)$ no dependen de que n sea par o impar; por otra parte, el extremo superior lo definimos como $(n-1)/2$ porque al ser n impar sólo estamos interesados en el recorrido $0 \le j \le (n-1)/2$), mientras que el valor a considerar sería $x = b - r = \frac{n-1}{2} - l'$, donde $l' = \frac{r-2}{2}$. Para el caso particular $p = q = 0'5$, el intervalo de la "*zona central*" es $\left(\frac{n}{2} - \frac{\sqrt{n}}{2}, \frac{n-1}{2}\right]$.

Para la segunda partida, $r = 2$, la situación es similar a lo que ocurre con la partida primera, ya que el valor asociado sería $(n-1)/2$, que siempre está en la "*zona central*"; aplicando la situación ii) del lema 2, obtenemos que al disminuir b, el valor de la segunda partida siempre debe aumentar, como así se observa en la tabla de Pascal.

Para la cuarta partida, $r = 4$, cuando $b = 6$, tendríamos que considerar el valor 2 de una binomial simétrica con $n = 7$, para la que la zona central es $(2'177, 3]$ y por tanto el valor está en "la cola", con lo que el valor de la partida cuarta debe disminuir al pasar a $b=5$, como de hecho se observa en la tabla. Si analizamos la situación $(4,5)$, veremos que $n=5$, por lo que la "zona central" es el intervalo $\left(1'382, 2\right]$; puesto que el valor asociado para este juego es $x=1$, vemos que de nuevo queda en "*la cola*", con lo que la probabilidad debe disminuir al pasar a $b=4$.

Conclusiones.

Una de las ideas principales que pretendemos remarcar en este apartado es la "*extrañeza*" que nos causa el hecho de que un problema planteado en la correspondencia entre Pascal y Fermat no haya sido abordado posteriormente por otros autores (al menos en lo que nosotros conocemos) a pesar de la gran cantidad de literatura existente dedicada a esta correspondencia.

Con respecto al problema en sí, destacar que el valor de la partida $r-ésima$ cuando a ambos jugadores les restaban b partidas es la probabilidad de un valor concreto $(b-r)$ en una distribución binomial con $2b-r-1$ ensayos, como se desprende inmediatamente de la fórmula aportada por Pascal. Esto nos conduce a que el comportamiento de los valores de la tabla de Pascal se deba a una propiedad de la distribución binomial, que nosotros hemos enunciado en el lema 2, que hace referencia al aumento o disminución de la probabilidad al pasar del valor j en la distribución X_n al valor $j-1$ en X_{n-2}, es decir, la base de la explicación del comportamiento de la tabla de Pascal es la comparación de las probabilidades $P\left[X_{n-2}=j-1\right]$ y $P\left[X_n=j\right]$, lo que finalmente se reduce al hecho de que el valor j de X_n se encuentre en lo que hemos denominado "*zona central*", "*cola*" o "*frontera*" de la distribución X_n.

Pascal sólo considera el caso de jugadores con igual habilidad. Nosotros hemos generalizado el resultado al caso de jugadores con distintas habilidades; para ello, es necesario introducir la hipótesis de que los jugadores apuestan una cantidad proporcional a su habilidad, pues en otro caso la valoración de las partidas ya no sería una generalización de la fórmula de Pascal.

Apéndice: Demostración del *Lema 1*.

Lema 1: $$W_A[(r,b)] = \binom{2b-r-1}{b-r} p^{b-r} q^{b-1}$$

Dmt.

Si suponemos que cada jugador apuesta una cantidad proporcional a su habilidad y consideramos que el total apostado es la unidad, entonces para el juego interrumpido (a,b) la ganancia esperada del jugador A es $q \cdot P_A\left[(a,b)\right] - p \cdot \left[1-P_A\left[(a,b)\right]\right] = P_A\left[(a,b)\right] - p$. Si ahora expresamos la ganancia esperada de la forma $\beta_A\left[(a,b)\right] \cdot q$, donde $\beta_A\left[(a,b)\right]$ es un valor del intervalo $\left[-1,1\right]$, entonces $P_A\left[(a,b)\right] = p + \beta_A\left[(a,b)\right] \cdot q$, es decir, la probabilidad de que el jugador A gane el juego es igual a lo apostado por dicho jugador más una parte proporcional a lo apostado por el jugador B. Cuando

$\beta_A\left[(a,b)\right]=0$, entonces cada jugador retira lo que apostó, cuando $\beta_A\left[(a,b)\right]>0$, entonces el jugador A gana $\beta_A\left[(a,b)\right]\cdot q$, mientras que si $\beta_A\left[(a,b)\right]<0$, entonces el jugador A pierde esa cantidad. Despejando de la última igualdad, obtenemos que $\beta_A\left[(a,b)\right]=\left(P_A\left[(a,b)\right]-p\right)/q$.

Recordemos que $W_A[(r,b)]$ es la diferencia entre la proporción que el primer jugador se lleva del otro al haber alcanzado el juego $(b-r,b)$ y esta misma cuando llegó al juego $(b-r+1,b)$, y por tanto

$$W_A\left[(r,b)\right]=\beta_A\left[(b-r,b)\right]-\beta_A\left[(b-r+1,b)\right].$$

Ahora, sustituyendo el valor de las proporciones β_A obtenemos:

$$W_A\left[(r,b)\right]=\frac{1}{q}\left(P_A\left[(b-r,b)\right]-P_A\left[(b-r+1,b)\right]\right)$$

y sustituyendo el valor de las probabilidades:

$$W_A\left[(r,b)\right]=\frac{1}{q}\left\{\sum_{i=b-r}^{2b-r-1}\binom{2b-r-1}{i}p^i q^{2b-r-1-i}-\sum_{i=b-r+1}^{2b-r}\binom{2b-r}{i}p^i q^{2b-r-i}\right\}.$$

Aplicado al factor segundo de la izquierda la igualdad siguiente,

$$\binom{2b-r}{i}=\binom{2b-r-1}{i-1}+\binom{2b-r-1}{i},$$

se obtiene la expresión,

$$\begin{aligned}\sum_{i=b-r+1}^{2b-r}\binom{2b-r}{i}p^i q^{2b-r-i} &= \sum_{i=b-r+1}^{2b-r-1}\binom{2b-r-1}{i-1}p^i q^{2b-r-i}+\sum_{i=b-r+1}^{2b-r-1}\binom{2b-r-1}{i}p^i q^{2b-r-i}+p^{2b-r}= \\ &=\sum_{i=b-r}^{2b-r-2}\binom{2b-r-1}{i}p^{i+1} q^{2b-r-(i+1)}+\sum_{i=b-r+1}^{2b-r-1}\binom{2b-r-1}{i}p^i q^{2b-r-i}+p^{2b-r}\end{aligned}.$$

Sustituyendo,

$$W_A\left[(r,b)\right]=\frac{1}{q}\left\{\sum_{i=b-r}^{2b-r-2}\binom{2b-r-1}{i}p^i q^{2b-r-1-i}(1-p)+p^{2b-r-1}-p^{2b-r}-\sum_{i=b-r+1}^{2b-r-1}\binom{2b-r-1}{i}p^i q^{2b-r-i}\right\}.$$

Operando,

$$W_A\left[(r,b)\right]=\left\{\sum_{i=b-r}^{2b-r-2}\binom{2b-r-1}{i}p^i q^{2b-r-1-i}+p^{2b-r-1}-\sum_{i=b-r+1}^{2b-r-1}\binom{2b-r-1}{i}p^i q^{2b-r-1-i}\right\}.$$

es decir,

$$W_A\left[(r,b)\right]=\left\{\sum_{i=b-r}^{2b-r-1}\binom{2b-r-1}{i}p^i q^{2b-r-1-i}-\sum_{i=b-r+1}^{2b-r-1}\binom{2b-r-1}{i}p^i q^{2b-r-1-i}\right\}$$

Finalmente,

$$W_A\left[(r,b)\right]=\binom{2b-r-1}{b-r}p^{b-r}q^{b-1}.$$

Bibliografía

BASULTO, J., CAMÚÑEZ, J.M. (2007,a), "Sobre una propiedad de la familia de Distribuciones Binomiales Simétricas detectada por Blaise Pascal (1654) en su resolución del problema de los puntos", *Estadística Española,* vol. 164, p. 33-58.

BASULTO, J., CAMÚÑEZ, J.M. (2007,b), *La geometría del azar. La correspondencia entre Pierre de Fermat y Blaise Pascal*, Ed. Nivola, Madrid.

BASULTO, J., CAMÚÑEZ, J. A., ORTEGA, F. J., PÉREZ, M. D. (2004). "Una fórmula casi mágica en la resolución de Pascal del problema de los Puntos.". Capítulo 9 de *Historia de la Probabilidad y de la Estadística (II)*, Editorial AC, Madrid, 157-169.

BERNOULLI, JACOB (1713), *The art of conjecturing*, Traducido por Edith Dudley Sylla, Jonhs Hopkins University Press, 2006, Baltimore.

DE MOIVRE, A. (1712), *De mensura sortis, seu, de Probabilitate Eventuun in Ludis a Casu Fortuito Pendentibus.* Traducido al inglés por B. McClintock en *Intern. Statist. Rev.*, 1984, 52, 237-262.

EDWARDS, A.W.F. (1987). *Pascal's arithmetical triangle*. Griffin, London.

HUYGENS, C. (1660) *Oeuvres Complètes*. 22 volúmenes. Société Hollandaise des Sciences. Nijhoff, La Haye. 1888-1950. Volumen usado: XIV.

KYRIACOPOULOS, L. (2000). "Peut-on tout de même parler d'un "triangle de Pascal"?". *Revue d'histoire des mathématiques,* 6, p. 167-217.

MONTMORT, P. R. (1713), *Essay d'Analyse sur les Jeux de Hazard.* Seconde Edition. Revûe et augmentée de plusieurs Leerte. Quillau, París. Reimpreso por Chelsea, New Cork, 1980.

PASCAL, B. (1963). *Oeuvres Complètes*, Edición de Lafuma, París.

Capítulo 25

Reproducción y muerte de la población mexicana: cálculos estadísticos y preceptos higiénicos a fines del siglo diecinueve

LAURA CHÁZARO GARCÍA
Estudios Avanzados del Instituto Politécnico de Zacatenco
México

Introducción

El discurso y prácticas higiénicos fueron algo más que un complemento de la clínica decimonónica del siglo diecinueve. Para uno de los más reconocidos higienistas de la Academia Nacional de Medicina de México (ANM), curar no sólo significaba devolver la normalidad al cuerpo, implicaba también "prolongar la duración de la vida y mejorar la condición física de la especie".[1] Con las estadísticas, un número importante de médicos vio un camino para hacer de la higiene un conocimiento y práctica capaz de evitar las enfermedades padecidas por la población. Efectivamente, preocupados por la dimensión colectiva de la enfermedad, los higienistas se preguntaban ¿cómo acceder a las causas de las enfermedades colectivas? Lo más importante era impedir la decadencia de lo que llamaban "raza mexicana", equivalente a la población.

A mediados del siglo diecinueve, la estadística médica consistía en recolectar y clasificar frecuencias de muertes y nacimientos, según enfermedades, los climas, sexos y costumbres y cálculos de mortalidad de distintas enfermedades. Con esos datos se pretendía ofrecer una explicación al supuesto lento crecimiento de la población, especialmente de la Ciudad de México, y encontraron una alta mortalidad, especialmente la infantil.

En este texto me pregunto por los significados de los cálculos y cifras de mortalidad que usaron los médicos de fines del siglo diecinueve para hablar de la "población mexicana".[2] El

---

[1] Liceaga, "Algunas", 1911, p. 4.
[2] Mis conclusiones se basan en los trabajos de médicos de la Academia Nacional de Medicina, miembros de la sección de higiene y estadística médica, grupo compuesto, entre los más influyentes, por José María Reyes, Gustavo Ruiz y Sandoval, Eduardo Liceaga, Demetrio Mejía, Luis Hidalgo y Carpio, Luis E. Ruiz, José G.

concepto de población resultante de recuentos estadísticos supuso valoraciones morales que los higienistas atribuyeron a los pobladores y al territorio nacional.[3] Esas valoraciones adquirían sentido según el "orden higiénico", es decir, las nociones médicas de normalidad y equilibrio moral. Los cálculos médicos implicaban valoraciones médicas; así los de mortalidad fueron referidos a la reproducción, valorada por los médicos como el ámbito de los cuidados maternos, del matrimonio y, en general, del goce de los cuerpos como fuentes de patologías y explicación última de la muerte.

Los estudios estadísticos de los higienistas no se limitaron a ser meros cálculos, adquirieron sentido y circularon cargados de preceptos higiénicos que, en rigor, estaban dirigidos a normar e intervenir la higiene pública y privada de la población. No se trata entonces de simples enunciados o juicios, sino de normas que el médico dotó con la autoridad de sus cálculos. Esto ilumina otro aspecto del concepto: si bien entre los valores estadísticos y los morales parece haber una continuidad explicativa (unos sustentan a los otros), la fuerza de las estadísticas, frente a los preceptos higiénicos, se fragmenta a tal grado que el diagnóstico higiénico, en materia de la población mexicana, aparece más como un mandato y sanción moral que como una descripción numérica.

Más que mudas cifras, los cálculos de la mortalidad y reproducción revelan ser explicaciones al tiempo que modos para controlar y sancionar las acciones de los sujetos que integraban aquella población. A imagen y semejanza de sus preceptos, los higienistas calcularon la mortalidad y ordenaron cifras tratando de conciliar su postura (política) que negó a la "raza mexicana" la capacidad de igualdad, de normalidad y civilización con la idea de que las poblaciones son un todo homogéneo (hecho de promedios), cuyas diferencias y desigualdades podían dar vida a un cuerpo nacional, sano y estándar.

La población: el cuerpo político del higienista

En México, a principios del siglo diecinueve, el concepto de población remitía a una diversidad de ideas. Una de ellas, esquematizando el discurso político de la época, era la noción de "pueblo" como "Estado-Nación".[4] Desde esa perspectiva el pueblo se pensaba como una abstracción, asociada a la representación ciudadana. Contigua a esta idea, la población era sinónimo de los "habitantes" que ocupan un territorio o una provincia; el conjunto de los miembros de un grupo racial, sin necesariamente pensarse en relación al gobierno o Estado.[5] Estas nociones de pueblo y población sin duda se alimentaron de una serie de prácticas y estrategias políticas de una elite preocupada por imaginar una nación moderna.[6] Según François-Xavier Guerra, en las sociedades hispanas, en el tránsito del antiguo régimen al moderno, las nociones de "nación", "pueblo" y "soberanía" se integraron al lenguaje político, aún cuando prevalecían relaciones de autoridad tradicionales.[7]

Lobato.

[3] Estas ideas las he formulado a través de múltiples lecturas, entre las más relevantes: Porter, "Making", 1994, pp. 389-407; Desrosières, *Politique*, 1993 y Desrosières, "¿cómo?", 1995.

[4] Una amplia gama de políticos de la época discutió las ideas de pueblo, nación y estado. Este texto no tiene el objetivo de discutirlas ni pretende ofrecer una definición de la noción de Estado nación.

[5] Bourget, *Déchiffrer*, 1989, pp. 287-289 ; Hobsbawm, *Naciones*, 1991 y Guerra y Lempérière *et al*., *Espacios*, 1988, pp. 130 y ss.

[6] Guerra, *México*, 1988, t. 1, pp. 190-194.

[7] Hobsbawm, *Naciones*, 1991, pp. 23-24, quien señala que sólo hasta fines del siglo diecinueve, la idea moderna del concepto de nación se definió como 'el conjunto de habitantes de un país regido por un mismo gobierno'. Antes podía referir a "colección de habitantes", aludiendo al territorio de donde se desciende, de la tierra de origen, territorio que sólo de manera fortuita podía corresponder a una unidad política.

Estas discusiones no fueron exclusivas de políticos, en ellas participaron los científicos, así como los médicos. Los higienistas se preocuparon por investigar y actuar sobre las causas físicas y morales que debilitaban a la población; tomada como un cuerpo susceptible de degenerar, enfermar ó vigorizarse. Múltiples datos dan a conocer sus patologías, el número de sus habitantes (sus muertes y nacimientos), el conjunto de sus características físicas (talla, peso, sexos) y sus tipos raciales, evaluaban sus costumbres y hábitos.

El uso y significado de las estadísticas médicas esta estrechamente ligado a las ideas médicas sobre las relaciones entre lo "físico" y lo "moral".[8] Inspirados en los médicos franceses conocidos como "revolucionarios" o "ideólogos" – entre ellos, Cabanis y Phillipe Pinel – los higienistas sostenían que muchos padecimientos están conectados o son consecuencia de hábitos, costumbres perniciosas y vicios. Un médico mexicano afirmó que las pasiones o los excesos pueden desencadenar padecimientos: "La mayor parte de los tísicos surgen de la lujuria y de la crápula; la gota y las enteritis de la glotonería y la de la intemperancia".[9] En esas premisas, la normalidad equivalía al orden político moral; así según el Dr. Rafael Lavista, médico legal y miembro de la ANM, la conexión entre las leyes de los organismos sociales y de los naturales se evidencia cuando los individuos, por incumplir o infraccionar sus obligaciones, "ocasiona(n) enfermedades sociales". Para él, el trabajo de "remediar las enfermedades del cuerpo social o del individuo en particular", trabajo propio de la "la jurisprudencia", sólo será efectivo cuando se atiendan las leyes de la fisiología.[10] Por eso, en materia de higiene y de medicina legal el médico no debía limitarse a conocer las leyes fisiológicas, debía además manejar las de los organismos sociales. Para aquella medicina, entonces, la normalidad dependía de un orden fisiológico y moral.

Basados en esos argumentos los higienistas pusieron en el mismo nivel hasta confundirlos las causas del desorden social y el origen de las patologías de la población. Las indagaciones estadísticas de los médicos permitían justamente explicar las causas de las enfermedades y, por extensión del desorden moral y político de la población. Los higienistas de la ANM la anunciaron como una herramienta que les permitiría hacer diagnósticos objetivos de las debilidades de la población y conducir a la población a un régimen de derechos y obligaciones, régimen tan abstracto como las cifras y los promedios estadísticos en los que se basaban. En otras palabras, las estadísticas aparecieron como la promesa para expresar con certeza científica un orden civilizado al cuerpo nacional.

La estadística médica calcula a la población mexicana

Los médicos mexicanos practicaron una estadística inspirada en los trabajos de los higienistas franceses como Louis-René Villermé[11] y otros miembros del grupo de los *Annales d'Hygiène Publique et Médicine Légale* (1829).[12] Y como ellos, decían seguir las ideas del astrónomo belga Adolphe Quetelet (1797-1874), autor del famoso *Sur l'Homme et le Développement de sus Facultés ou Essai de Physique Sociale* (1835).[13] La influencia del belga en México fue

[8] Williams, *Physical*, 1994, pp. 45 y 52.
[9] Malanco, "Fisiología", 1897, pp. 408-409.
[10] Lavista, "Relaciones", 1895, pp.8 y 6.
[11] Sobre los higienistas franceses véase, entre otros, Coleman, *Death*, 1982, pp. 225-226 y 230 y La Berge, *Mission*, 1992, p.17 y ss. De Villermé. Por ejemplo, "Hygiène", 1829, t. 1, pp.1-100.
[12] Los autores de los *Annales* fueron miembros del Conséil de Salubrité francés; médicos y burócratas comprometidos con las reformas de salud pública, entre ellos: Gabriel Andral, profesor de Higiene (1828), Esquirol, Orfila, Parent-Duchatelet y Louis-René Villermé. Ackerknecht, "Hygiene", 1948; La Berge, "Early", 1984.
[13] Quetelet, *Sur l'homme*, 1991. Sobre Quetelet y sus nociones de estadística moral: Stigler, *History*, 1986, pp.

amplia, alcanzó también a los estadísticos e ingenieros[14]; la mayoría lo siguió bajo una interpretación higienista: el cálculo de las leyes de población servía para propiciar la vigorización fisiológica y económica del país. Bajo estas premisas por estadística médica se entendía acumular frecuencias y calcular promedios y proporciones.[15] Se trataba de cálculos que servían para determinar qué eventos provocan más muertes e identificar la cifra de la *reproducción normal* de la población. Uno de los propósitos de la estadística médica era identificar ¿cuáles son las causas de las patologías colectivas?

Las estadísticas médicas trataban con las "mayorías",[16] es decir, investigaban en una serie larga de datos lo característico o común a un grupo de individuos; las mayorías revelan, aunque de forma abstracta, las características del conjunto. Entendida así, la "población" se convertía en una entidad *homogénea* a la que se le podían atribuir *regularidades*. Aunque compuesta de la suma de individuos, las poblaciones fueron pensadas como las "mayorías", apreciables por medio de las probabilidades estadísticas.

Los médicos pretendieron calcular las leyes de la población mexicana valiéndose de dos tipos de promedios: el llamado "cálculo de la Mortalidad media" (*Mm*) y "la Vida media", también llamada "mediana de la vida" (*Vm*).[17] La mortalidad media, también conocida como "promedio anual aritmético del número de muertes de una población", es una proporción del tipo[18]:

$$Mm = \mathrm{d/a}$$

Donde "d" es igual el total de defunciones de un periodo definido y "a" el número de años de ese periodo. Para los médicos, la *mortalidad media* mide el desgaste o envejecimiento natural de la población, es decir, sus pérdidas constantes y necesarias. Con este tipo de cálculos, el médico decía determinar cuánto una enfermedad, según el número promedio de muertes que causó, contribuyó a la mortalidad total de una población. Así definía si la mortalidad de una población estaba dentro de los límites naturales de envejecimiento o si respondía a causas patológicas o degenerativas. El Dr. Gustavo Ruiz y Sandoval decía a propósito de ello que: la mortalidad de un pueblo, bien computada y apreciada (...) indica su adelanto o atraso en todos los ramos de la higiene pública y privada. Y a nadie se ocultará que las tablas de mortalidad de muchos años continuados, enseñan, con el aumento o disminución de su cifra, los pasos que una sociedad da para acercarse o alejarse de su bienestar.[19]

El promedio de la vida, o la vida media de la población indicaba la fortaleza o debilidad de los habitantes. En palabras del Dr. Ruiz y Sandoval, la *vida media* expresa "el término medio de los años de vida de los habitantes de una población"[20], es decir, indica cuántos

---

161-182; Porter, *Rise*, 1986, pp. 40-55, entre los más clásicos.

[14] Antonio García Cubas, miembro de la Sociedad Mexicana de Geografía y Estadística habla de esa influencia en "Sesión en honor de Quetelet", *El siglo XIX,* 28 de julio de 1874, p. 1.

[15] Las características de la estadística médica las he abordado también en: Cházaro, "Midiendo el cuerpo de una nación. Ensayo sobre la estadística médica en México a finales del siglo XIX", tesis doctoral en Filosofía, Facultad de Filosofía y Letras, UNAM, México, 2000.

[16] Las mayorías o maximums siempre implicaban al opuesto, minoría o minimums. El *Maximum*, según una definición de la época es "un valor más grande y el *minimum* un valor más pequeño que los valores que preceden y los que siguen dentro de ciertos límites", Bouillet, *Dictionnaire*, 1874, p. 1053.

[17] Este término hoy se conoce como "esperanza de vida" y refiere a poblaciones "estacionarias". El cálculo de la duración de la vida media apareció a mediados del siglo XVIII. Hay que notar que es esta medida es distinta de la *vida mediana*, pero también conocida entre los probabilistas como *vida media*. Entre los médicos mexicanos, ambos términos se confundían. Dupaquier, *Invention*, 1996, pp. 31-32.

[18] Esta es mi notación; en los textos médicos los cálculos no son visibles; generalmente los médicos no acostumbraban a presentar sus procedimientos aritméticos.

[19] Ruiz y Sandoval, *Estadística*, 1872, p. 11.

[20] El cálculo y sus bases están explicados en Ruiz y Sandoval, *Estadística*, p. 9 y en Ruiz y Sandoval y Quintas

habitantes sobreviven a la mortalidad patológica de una población. Los higienistas mexicanos la obtenían así:

$$Vm = p/d$$

Donde "p" es igual a la población viva y "d" el total de defunciones de esa población en un periodo determinado de años. Con estos promedios los médicos buscaron interrogar las cifras de mortalidad y natalidad, siguiendo las tesis estadísticas de Adolphe Quetelet, especialmente la referida a la *estabilidad de las poblaciones*[21].

En *Sur l'homme* Quetelet caracterizó a las poblaciones como entidades que tienden a estabilizar el número de nacimientos y muertes; así aunque se observaran índices mayores o menores de esterilidad, siempre se registraría un número de nacimientos "sorprendentemente" *estable*. Esto se debía, según él, a que siempre habría parejas fértiles que compensaran las carencias en el número de nacimientos.[22] De hecho, Quetelet postuló que un crecimiento normal se asevera cuando hay un equilibrio entre el número de muertes y de nacimientos. Es decir, a mayor número de muertes, la población responde con más nacimientos y a mayor número de nacimientos más muertes; predominando el equilibrio entre ambas tendencias: el número de nacimientos es menos grande cuando el número de decesos es igualmente el menos fuerte: lo que coincide muy bien con el señalamiento de Malthus de que el número de nacimientos aumenta cuando se ha hecho un vacío en la población, aún después de flujos destructores. Podemos creer que esta coincidencia se debe a que la mortalidad, que es muy grande entre los niños, crece en razón de los nacimientos.[23]

Los higienistas y el propio Quetelet no creían en la tesis de un crecimiento geométrico de las poblaciones, evidentemente no coincidían con las ideas del inglés Thomas R. Malthus (1766-1834).[24] Una enorme cantidad de obstáculos frenaban (regulando) el crecimiento poblacional, por ejemplo, la edad de los matrimonios, el efecto de las estaciones y el nivel de las subsistencias del país. Esos factores provocaban que las poblaciones tendieran "de más en más, a hacerse estacionarias", puesto que "la población encuentra, en su tendencia a crecer, las causas que previenen funestas catástrofes, resultantes de un demasiado pleno, si puedo expresarme así".[25] En desacuerdo con Malthus, Quetelet definió una ley compensatoria que impedía que las poblaciones sufrieran incrementos o decrementos bruscos: "los países donde hay más hijos por matrimonios, son también aquellos que tienen una gran mortalidad"; a menor probabilidad de vivir, mayor fecundidad.[26] Por eso, para él, la reproducción de las poblaciones dependía más del cuidado de los vivos que de nuevos nacimientos: "la prosperidad de los estados (...) debe consistir menos en la multiplicación que en la conservación de los individuos que los componen".[27]

Arroyo, "Estadística médica", 1874, p. 5.

[21] Stigler, *History*, 1986, pp. 165-166.

[22] Este argumento no fue exclusivo de Quetelet; antes lo defendió el ilustrado Montyon Véase *Recherches*, 1774, pp. 34-35, obra que fue publicada bajo el nombre de Moheau pero que actualmente se le atribuye a Montyon.

[23] Quetelet, "Mémoire", s/d. p. 500.

[24] No es mi propósito comparar las ideas de Malthus y Quetelet, pero sí subrayar que a pesar de sus diferencias coincidían, por ejemplo, el peso que le dieron al factor de las subsistencias con relación a los ritmos de crecimiento de la población. Y si bien es cierto que para el belga no tenía sentido proponer la posposición del matrimonio o la "abstención moral" como un *obstáculo preventivo* al crecimiento de las poblaciones, él igual que los médicos mexicanos, subrayó la necesidad de controlar y circunscribir el ejercicio de la sexualidad al matrimonio. Malthus, *Ensayo*, 1960, pp. 7-12 y 14-19.

[25] Quetelet, "Mémoire", s/d, p. 501.

[26] Quetelet, *Sur l'homme*, 1991, pp. 79-80.

[27] *Ibid.*, p. 145.

Convencidos de las ideas de Quetelet, los higienistas creían que para determinar el tipo de crecimiento de la población, les bastaba con calcular la mortalidad y la natalidad. El doctor José María Reyes, uno de los higienistas más reconocido de la época, enfatizó que "hoy está plenamente averiguado que más que por la fecundidad de las mujeres, las poblaciones crecen por la disminución de los fallecimientos de los niños; ellos dan el factor principal en el aumento de la vida media, y si la natalidad suministra su contingente, éste se halla subordinado al anterior".[28] Esta opinión, en principio aritmética y metodológica, tuvo importantes consecuencias en la forma como los médicos pensaron a la población mexicana y sus problemas para reproducirse.

Las cifras de la mortalidad frente al problema de la reproducción de la población

Para mediados del siglo diecinueve, los higienistas habían escrito un gran número de estudios estadísticos y habían hecho importantes hallazgos sobre la mortalidad media, en especial en la Ciudad de México. José Ma. Reyes, Gustavo Ruiz y Sandoval y Demetrio Mejía, entre otros higienistas de la ANM, encontraron que entre 1833 y 1874, en promedio por año, hubo una mortalidad de 8841.6 individuos.[29] Retomando las estimaciones del Dr. Ruíz y Sandoval, esos médicos coincidieron al señalar que en ochenta años la población de la ciudad de México registró un aumento de sólo 112,074 habitantes, lo que reforzó la idea de que la población crecía muy lentamente. Pero lo más alarmante fue hallar que del total de la mortalidad media anual de la ciudad de México, el 45% correspondía a la población infantil. La mortalidad infantil se convirtió en la principal causa de la despoblación del país, el obstáculo contra la reproducción de la población.

Con fragmentarios censos del país y de la Ciudad de México, los higienistas rechazaron las predicciones de un supuesto acelerado crecimiento de la población. Se apegaron a una tesis poblacionista que creía que controlando la mortalidad se incidiría en el número de individuos de la población. Es decir, el número y vigor de la población no resulta ni del incremento de los nacimientos, tampoco –como defendían algunos políticos de la época – incentivando la inmigración de europeos. El Dr. Reyes fue claro: "en vano es que se proclame la colonización como base indispensable de mejora social, si no se cuida del aumento de la población en su edad más delicada".[30] Como Reyes, el Dr. Ruiz y Sandoval hizo eco de esa conclusión: "antes que procurar la inmigración, (el Estado) debe ver que no mueran tan temprano los individuos cuya custodia y bienestar le están encomendados".[31] El hallazgo estadístico del excesivo número de muertes infantiles exigía explicaciones. José María Reyes opinó que la mortalidad infantil encuentra explicación, entre otras razones, en los matrimonios ilegítimos y en la inmoralidad de los hombres y las mujeres pobres: La miseria, la vagancia, la embriaguez y la inmoralidad están muchos más generalizadas, y se sabe bien que estos vicios obran como principales agentes en la seducción, en el abandono de las madres y los hijos y en el aumento de la prostitución bajo todas sus formas.[32]

Pronto el énfasis fue puesto en materias de higiene moral y no en los cálculos. La reproducción fue asociada al ámbito del matrimonio, los cuidados maternos y, en general, los excesos y gozos de hombres y mujeres. Los médicos prácticamente no sistematizaron datos

---

[28] Reyes, "Higiene", 1878, p. 384.

[29] Es interesante constatar que al menos estos tres médicos, basados en datos distintos, encontraron promedios de mortalidad para la Ciudad de México muy cercanos. Ver Mejía, "Estadística", 1879, p. 284.

[30]* Reyes, "Memoria", 1867, p. 174.

[31] Ruiz y Sandoval, *Estadística*, 1872, p. 18.

[32] Reyes "Higiene", 1878, p. 379.

estadísticos de matrimonialidad, número de matrimonios, la fertilidad, el número de hijos por mujeres, o la fecundidad de la población, el número de hijos por matrimonios. Si para Adolphe Quetelet, el cálculo de las curvas de crecimiento de la población exigía considerar ese tipo de variables; ¿por qué los médicos no los hacían? No fue por desinterés, ni por falta de información estadística;[33] más bien, porque las interrogantes estadísticas parecían tener respuestas en la vida moral y pública, no en los cálculos. Efectivamente, las prácticas asociadas a la reproducción, no eran entidades de cálculo, sino de control e intervención moral. De ahí el interés médico por ordenar, con normas y prácticas médicas, la vida reproductiva de los habitantes del país. Puesto de otro modo, el significante de los cálculos de la mortalidad no fue la proporción "d/p", sino la vida moral de los pobladores, especialmente, lo relativo a sus prácticas sexuales, el motivo mismo de la existencia de la población. Sustentados en la autoridad de los números, la atención se centró en los gozos, los hábitos sexuales y, en la vida moral de la población, como las fuentes de la degeneración y muerte.

Los higienistas exploraron la "mortalidad" relacionándola a la degeneración de las razas, la vida sexual de los hombres y las mujeres, las uniones ilegales o la antihigiene. Estos análisis, basados en preceptos y juicios médicos tomaron la fuerza de las conclusiones numéricas. Entre los cálculos de la mortalidad y los juicios sobre la reproducción se revela la ambigüedad del argumento higienista, hecho de los valores numéricos y la sanción moral.

La degeneración y los gozos generan debilidad y muerte

Los médicos de la época no confiaban en la capacidad "natural" de la población mexicana de preservarse, no sin su intervención. Distintas tesis los inspiraron, entre otras, la convicción de que las variaciones físicas y morales observadas entre los individuos pueden ser signos de una supuesta degeneración de las razas y los resultados de los estudios sobre la variación biológica y la herencia.[34]

Una gran mayoría de médicos mexicanos coincidió en consignar la acción de la "herencia mórbida" como un posible mecanismo de degeneración de la población. Tomada como segura la hipótesis de que por vía hereditaria las patologías merman a los individuos, los higienistas afirmaban que ciertos padecimientos, especialmente los epidémicos, podían transmitirse a la prole como factores predisponentes a la enfermedad. Así el doctor Román Ramírez, médico legista de la ANM, decía que ciertas enfermedades inoculaban a su descendencia con "gérmenes debilitantes".[35] El ejemplo más socorrido eran los hijos de alcohólicos, a quienes se les consideraba predispuestos "a la sordera y a las afecciones del sistema nervioso, convulsiones, epilepsia, alienación mental, a la depravación de los instintos, a los impulsos fatales e irresistibles".[36] Sin contar con cifras concluyentes, los médicos pensaban que reduciendo los descendientes de los afectados por enfermedades llamadas de familia, entonces disminuiría la mortalidad infantil. Como si se tratara de una conclusión lógica, los interesados en la mortalidad centraron la mirada en la naturaleza de las uniones conyugales.

[33] Para un análisis de este tema en la Europa del siglo XIX, véase La Vergata, 2000, pp. 191-192.

[34] No me extenderé aquí en la tesis de la "debilidad" o "degeneración" racial que, sin duda, fue central a la higiene mexicana del siglo diecinueve. Sobre el tema, Pick, *Faces*, 1993; Harris, *Murders*, 1991, p. 57. Para el caso mexicano, véase Urías Horcasitas, *Indígena*, 2000, p. 76 y ss.

[35] Ramírez, "Nociones", 1901, p. 69. Las investigaciones médicas sobre los mecanismos hereditarios le dieron un papel dominante a las patologías. Padecer enfermedades "crónicas" se afirmaba, frecuentemente, era propicio para procrear hijos de débil constitución o propensos a otras enfermedades, progresivas, ideopáticas o relacionadas con la idiosincrasia individual. Obsérvese que estos atributos individuales, según los médicos, se extienden a las poblaciones. López Beltrán, "Maladies", 1995, pp. 340-341.

[36] Kulburn, *Cours*, 1867, p. 384.

Buscando “neutralizar los elementos mórbidos, que cada uno de los futuros aportará”,[37] los higienistas crearon una rudimentaria “teoría” que confundió la sexualidad y las leyes de la reproducción de la población. Como si se tratara de lo mismo, hicieron de la sexualidad un fenómeno del crecimiento de las poblaciones; en ello fue implícito el gozo de los cuerpos, que se le incluyó al ámbito del desequilibrio y del vicio. La normalidad en el ámbito de la sexualidad equivale entonces a los límites legales del matrimonio; todo lo que sucediera fuera del matrimonio iba contra las leyes de fisiológicas y jurídicas, atentaba contra el crecimiento de la población. Así entendido, para los higienistas sólo interviniendo y limitando el goce de los cuerpos, se aseguraría el crecimiento de la población.[38]

El médico con sus estadísticas y el legislador con sus leyes buscaron regular la vida y las relaciones entre los sexos. En su monumental Compendio de medicina legal (1877), los doctores Luis Hidalgo y Carpio y Gustavo Ruiz y Sandoval evaluaron, entre otros temas, lo relativo al matrimonio y se mostraron de acuerdo con el Código Civil cuando ordenó: “la procreación es el fin esencial del matrimonio”; sólo en el matrimonio las uniones entre hombres y mujeres tienen el derecho legítimo de “perpetuar a la especie”, el placer de reproducirse.[39] Higienistas y jueces pretendieron mostrar que el matrimonio frenaban las razones de la mortalidad y despoblamiento.[40]

Otras variables afectaban los promedios de mortalidad: “la salud y robustez de los padres, la influencia del parentesco, las faenas y los trabajos de la mujer durante la preñez, las miserias y las afecciones morales”.[41] Igualmente, múltiples detalles ligados a los cuidados del niño fueron objeto de la mirada médica en búsqueda de redondear los resultados estadísticos de la mortalidad. La “ignorancia femenina” en lo relativo a la higiene y a la nutrición también se volvió causa de la mortalidad infantil: “Sin numeración precisa” pero porque lo “hemos observado en México”, el doctor Reyes afirmó que las madres, o bien sustituyen la leche materna “con atoles” o “con una nodriza cuya leche no se analiza debidamente”. En todos los casos, aseguró el médico, se transmite el “germen de enfermedades, especialmente la enteritis”.[42]

En este esquema, las soluciones no vendrían del control de la mortalidad sino de los gozos dentro de los límites del hogar y de los vicios antihigiénicos; de la intervención médica a la vida de los pobres, consideradas presas fáciles de la muerte. Para Ruiz el alto porcentaje de niños muertos no podía tener otra causa que la “miseria” de los padres: al “entregarse a toda clase de excesos”, dan “hijos débiles y que a la menor causa les hará contribuir a la mortalidad en este periodo”.[43]

A pesar de la prolijidad en detalles, los conteos acerca de la mortalidad, frecuentemente, confundían la categoría de “pobre” con la de “indio”. En las estadísticas el pobre se volvió intercambiable con la de población indígena, ambas contaban como los condenados a la

[37] *Ibid*., p. 385.
[38] Frida Gorbach llama la atención sobre esta característica del pensamiento médico, especialmente de la medicina legal en que se distingue el goce del placer, entendiendo este último como acto reglamentado jurídicamente. Véase Gorbach, “Placer”, 1998, pp. 7-19. Lo que interesa subrayar acá es ese aspecto de la medicina: devolver al cuerpo al estado normal, es decir, ponerlo en un estado de ausencia del dolor pero también del gozo absoluto.
[39] Hidalgo y Carpio y Ruiz y Sandoval, *Compendio*, 1877, t. 1, pp. 71 y 74.
[40] Nicholson, “Sex”, 1992, pp. 419-420. El pronatalismo de los médicos mexicanos no impidió que tuvieran una concepción casi similar a la de T.R. Malthus en lo que se refiere al papel de la sexualidad en la reproducción de la población. Así, fueron partidarios de controlar la sexualidad para evitar la degeneración de la raza mexicana.
[41] Reyes “Higiene”, 1878, p. 379.
[42] *Ibid*., p. 383.
[43] *Ibid*.

muerte, a la degeneración racial, a la criminalidad, al alcoholismo y a la prostitución. Valiéndose del censo de 1895, levantado por la Dirección General de Estadística, el doctor José G. Lobato concluyó que "Con la raza indígena en nuestro país, sucede lo que con la negra en Estados Unidos: que la mortalidad es mayor" (…) por la falta de civilización".[44] Para los médicos, los indígenas del siglo diecinueve habían perdido el resplandor de generaciones pasadas, "envilecidos por enfermedades epidémicas como el tifo, la fiebre amarilla o las diarreas".[45]

Los análisis higiénicos sobre la población no fueron mera enumeración y cálculos estadísticos, cargaban juicios y valoraciones médicas sobre la población mexicana. Esto no significa que la estadística médica fuera una mera artimaña. La noción de población surgió justamente del debate entre los valores estadísticos y los valores políticos que pretendían ordenar a la sociedad mexicana de la época. Este aspecto queda claro con la distancia que se generó entre la población ideal pensada a partir de las estadísticas y la población ideada por sus propios juicios y preceptos higiénicos. La noción de población surgida de las estadísticas médicas permite ver sus límites, escindido entre la autoridad del conocimiento estadístico y sus valoraciones morales.

Igualdad aritmética y política versus los preceptos higiénicos

La noción médica de la población está a la base del diagnóstico político acerca de la nación mexicana. No es raro encontrar estudios médicos que usaron metáforas políticas, como tampoco fueron escasos los políticos que defendieron las tesis de los higienistas.

Un ejemplo de ello fueron las ideas del influyente abogado y educador Justo Sierra[46] quien, inspirado en el inglés Herbert Spencer, opinó que la población mexicana aún no alcanzaba el estatus de "sociedad" pues seguía desgarrada y dividida entre razas y costumbres heterogéneas, sumida en el desorden y la enfermedad. Según esta lectura, las poblaciones, constituida de diferentes razas y con distintos grados de civilización tienden, por naturaleza, a degenerar, provocando debilidad racial, despoblamiento y una mortalidad excesiva.[47] Para Sierra las soluciones políticas al problema no podían derivar de procurar igualdad entre los sujetos de la población sino de un amplio poder. Los higienistas coincidían con Sierra en que la población es una entidad propensa a la mortalidad, pero, a diferencia de él, buscaron eliminar las causas de sus patologías a partir del diagnóstico estadístico. Los médicos esperaban intervenir la vida pública y privada de la población esperando borrar las causas de sus patologías y mortalidad. Así, desde la abstracta igualdad de las *mayorías*, es decir, la igualdad que resultaba de promediar sus diferencias, esperaban propiciar el crecimiento de una población como una *entidad homogénea*, sujeta a regularidades, predecible y manejable. Aunque los médicos siempre subrayaron las diferencias morales y fisiológicas de los habitantes del país, sus estadísticas los representaron en categorías (razas, edades, sexos, muertes/enfermedades) que los integraban a un todo contable, que podían ser expuestas en cuadros legibles y regulares. En otras palabras, la fuerza de la operación higiénica residía en la posibilidad de crear una población como una entidad homogénea, como si se tratara de manipular promedios.

---

[44] *Lobato, "Higiene"*, p. 33. Ver también, Urías Horcasitas, *Indígena*, 2000, pp. 61-66 .
[45] Otero, "Geografía", 1894, pp. 61-114.
[46] Justo Sierra fue uno de los críticos de la Constitución liberal de 1857 pues decía que la sociedad mexicana no estaba preparada para la democracia postulada por los principios constitucionales. Sierra, *Obras*, 1986, p. 282.
[47] Urías señala que la obra de Sierra y de los historiadores liberales de su época coincidía en que para reconstruir a la sociedad mexicana sumida en la anarquía no debía perder de vista las "contradicciones existentes entre los principios constitucionales y las condiciones históricas que determinaban el grado de evolución de la sociedad", *Historia*, 1996, p. 173.

Sin embargo, frente a esa operación higiénica basada en las estadísticas, permanecía la enfermedad y la pobreza; los grandes números de la mortalidad infantil. En el discurso por momentos es visible la distancia entre el supuesto de homogeneidad estadística y la vivencia de la confusión que genera la enfermedad y la diferencia. Efectivamente, el abstracto equilibrio de los nacimientos y las muertes postulado por los principios estadísticos, parece más un ideal ó abstracción. Paradójicamente, una de las misiones de la higiene había sido proveer a la población, a partir del frío diagnóstico de las cifras estadísticas, de normas "justas" y "objetivas", capaces de regenerar y unificar a la población. Sin embargo, frente a la recurrente enfermedad y la muerte, esas estadísticas y sus promedios parecen perderse, como meras abstracciones ideales. Precisamente, el concepto médico de la población está envuelto en esa paradoja: por un lado, pretende ser el diagnóstico frío hecho de la igualdad abstracta de los números, pero, para intervenir a los cuerpos, opta por convertirse en mera recomendación moral, mero precepto, desnudo de estadísticas y cálculos. Entonces, ¿dónde quedaron las leyes de la estadística y sus cálculos?

Como parte de los preceptos higiénicos, los cálculos estadísticos adquieren sentido y fuerza no como herramientas descriptivas sino como fuentes de normas y orden. Para determinar las leyes de la población, los médicos recurrieron a proporciones como la de "d/a", sus resultados servían, más allá de describir tendencias, para extraer un orden moral. Así, las estadísticas médicas, de postularse como diagnósticos aritméticos, en los textos higiénicos se convirtieron en recomendación moral sobre los marginados con la pretensión de detener la inmoralidad y los goces. El diagnóstico de las estadísticas de la población mexicana, sancionada como antihigiénica, degenerativa y de vida breve, se pierde ante la sanción moral.

Bibliografía

ACKERKNECHT, ERWIN.: "Hygiene in France, 1815-1848", *Bulletin of the History of Medicine*, vol. XXII, 1948.

BOULLET, M.N. : *Dictionnaire Universel des Sciences, des Lettres et des Arts*, Paris, Hachette, 1874.

BOURGET, MARIE NOELLE. : *Déchiffrer la France. La Statistique Départamentale à l'Epoque Napolénoienne*, Paris, Editions des Archives Contemporains, 1989.

COLEMAN, WILLIAM.: *Death is a Social Disease. Public Health and Political Economy in Early Industrial France*, The University of Wisconsin Press, 1982.

CHÁZARO, LAURA.: *Midiendo el cuerpo de una nación. Ensayo sobre la estadística médica en México a finales del siglo XIX*, tesis doctoral en Filosofía, Facultad de Filosofía y Letras, UNAM, México, 2000.

DESROSIERES, ALAIN. : *La Politique des Grands Nombres. Histoire du Raison Statistique*, Paris, Editions La Découverte, 1993.

DESROSIÈRES, ALAIN. : "¿Cómo fabricar cosas que se sostienen entre sí? Las ciencias sociales, la estadística y el Estado", *Archipiélago. Cuaderno de crítica de la cultura. Primera Carpeta*, Madrid, núm. 20, 1995.

DUPAQUIER, JACQUES.: *L'Invention de la Table de Mortalité*, Paris, PUF, 1996

ESCALANTE GONZALBO, FERNANDO.: *Ciudadanos imaginarios*, México, El Colegio de México, 1992.

GORBACH, FRIDA.: "El placer del encierro: La imagen de un hermafrodita", *Argumentos. Estudios críticos de la sociedad*, núm. 29, UAM-X, abril de 1998.

GUERRA, FRANCOIS, XAVIER. Y A. LEMPÉRIERE ET AL.: *Los espacios públicos en Iberoamérica. Ambigüedades y problemas, siglos XVII-XIX*, México, FCE, 1988.

GUERRA, FRANCOIS XAVIER.: *MÉXICO: Del antiguo régimen a la revolución*, México, FCE, t. 1, 1988.

HACKING, IAN.: *La domesticación del azar*, Madrid, Gedisa, 1991.

HARRIS, RUTH.: *Murders and Madness: Medicine, Law and Society in the Fin de Siècle*, Oxford, Clarendon Press, 1991.

HIDALGO Y CARPIO, LUIS Y GUSTAVO RUIZ Y SANDOVAL.: *Compendio de medicina legal arreglado a la legislación del Distrito Federal*, t. I, México, Imp. de Ignacio Escalante, 1877.

HOBSBAWM, ERIC.: *Naciones y nacionalismo desde 1780*, Barcelona, Crítica, 1991.

LA BERGE, ANN F., *Mission and Method. The Early Nineteenth-Century-French Public Health Movement*, Cambridge, Cambridge University Press, 1992.

LA BERGE, ANN F.: "The Early Nineteenth Century French Public Health Movement: The Disciplinary Development and Institutionalization of Hygiene Publique", *Bulletin of the History of Medicine*, vol. 58, numb. 3, 1984.

LA VERGATA, ANTONELLO.: "Biology and Sociology of fertility. Reactions to the Malthusian threat. 1798-1933", en Brian Dolan (edit.), *Malthus, Medicine and Morality. 'Malthusianism' after 1798*. Clio Medica 59, Ámsterdam, Rodopi, 2000, pp. 189-210.

LAVISTA, RAFAEL.: "Relaciones entre la medicina y la jurisprudencia", *Concurso Científico*, México, Oficina Tipográfica de Fomento, 1895.

LICEAGA, EDUARDO.: "Algunas consideraciones acerca de la higiene social en México", *Concurso científico y artístico del Centenario, Academia Mexicana de Jurisprudencia y Legislación*, México. Tip. Vda. de F. Díaz de León, Sucs., 1911.

LIER, E.: "Ginecología. La esterilidad en los matrimonios", *Gaceta Médica de México*, tomo XXV, núm. 12, 1890, pp. 221-240.

LOBATO, JOSÉ GUADALUPE.: "Higiene. Sociología en sus relaciones con la demografía y demología mexicanas", *Gaceta Médica de México*, tomo XV, núm. 16, 1880, pp. 357-371.

LÓPEZ BELTRÁN, CARLO.: "'Les maladies héréditaires' 18th Century Disputes in France", *Revue of the History of Science*, XVII, 3, 1995.

MALANCO, FERNANDO.: "Fisiología psicológica. Conexión entre lo físico y lo moral del hombre. Ventajas que de ella puede sacar la medicina", *Gaceta Médica de México*, tomo XXXIV, núm. 15, 1 de agosto de 1897, pp. 406-411.

MEJÍA, DEMETRIO.: "Estadística de la mortalidad en México", *Gaceta Médica de México*, tomo XIV, núm. 14, 15 de junio de 1879, pp. 273-301.

MALTHUS, THOMAS ROBERT.: *El ensayo sobre la población*, México, Fondo de Cultura Económica, 1960.

MOHEAU. : *Recherches et Considérations sur la Population de la France*, París, Moutard, Imprimeur de la Reine, 1774.

NICHOLSON, MERVYN. : "Sex and Spirit in Wollstonecraft and Malthus", *Journal of the History of Ideas*, Vol. LI, 1990, pp. 401-421.

ORVAÑANOS, DOMINGO. : "Higiene pública. Algunas consideraciones sobre la mortalidad en la República Mexicana", *Gaceta Médica de México,* tomo XXXVI, núm. 2, 15 de enero de 1899, pp. 22-36.

OTERO, MANUEL.: "Geografía médica. Apuntes para el estudio del clima y enfermedades propias a la ciudad de San Luis Potosí", *Gaceta Médica de México*, tomo XXXI, núm. 4, 1894, pp. 61-114.

PICK, DANIEL.: *Faces of Degeneration. A European Disorder, c. 1848- c.*1918, Cambridge, Cambridge University Press, 1993.

PORTER, THEODORE.: "Making Things Quantitative", *Science in Context*, Cambridge, vol. 7, number 3, Autumn 1994, pp. 389-407.

PORTER, THEODORE.: *The Rise of Statistical Thinking. 1820-1900*, Princeton University Press, 1986.

QUETELET, ADOLPHE. : "Mémoire sur les Lois des Naissances et de la Mortalité à Bruxelles", S.l., *Extrait des Mémoires de l'Académie Royale des Sciences et Belles Lettres de Bruxelles*, t. III, s/d.

QUETELET, ADOLPHE. : *Sur l'homme et les développements de ses facultés. Essai de Physique Sociale*, Paris, Fayard, 1991.

QUINTAS ARROYO, Juan.: "Estadística médica", GMM, 1874.

RAMÍREZ, ROMÁN.: "Nociones preliminares de patología", *Resumen de la Medicina legal y ciencias conexas*, México, Oficina Tipográfica de Fomento, 1901.

REYES, JOSÉ MARÍA.: "Estadística de mortalidad en la capital, con arreglo al censo de su población. Su estado patológico. Primera y segunda parte", *Boletín de la SMGyE*, México, Imprenta del Gobierno, 1869.

REYES, J.M.: "Higiene pública. Mortalidad de la niñez", *Gaceta Médica de México*, tomo XIII, núm. 20, 11 de julio de 1878, pp. 377-385.

REYES, J. M.: "Memoria leída por el secretario del Consejo Central de Salubridad el día 17 de enero de 1867, relativa a los trabajos de esta corporación en el año próximo pasado", Documento del Fondo Reservado, Lafragua, UNAM**,** 1867.

RUIZ Y SANDOVAL, G.: *Estadística de la mortalidad y sus relaciones con la Higiene y la patología de la Capital*, México, 1872.

SIERRA, JUSTO.: *Obras Completas. Evolución política del pueblo mexicano*, México, UNAM, 1986.

STIGLER, STEPHEN.: *The History of Statistics. The Measurement of Uncertainty before 1900*, Belknap Press of Harvard University Press, 1986.

URÍAS HORCASITAS, BEATRIZ.: *Indígena y criminal. Interpretaciones del derecho y la antropología en México. 1871-1921*, México, Universidad Iberoamericana, 2000.

URÍAS HORCASITAS, BEATRIZ.: *Historia de una negación: la idea de igualdad en el pensamiento político mexicano del siglo XIX*, México, UNAM, 1996.

VILLERME, LOUIS-RENE. : "Hygiène Publique. Mémoire sur la Mortalité dans les prisons", *Annales d'Hygiène Publique et de Médicine Légale*, t. I, Paris, 1829.

WILLIAMS, ELIZABETH.: The Physical and the Moral. Anthropology, Physiology, and Philosophical Medicine in France, 1750-1850, Cambridge, Cambridge University Press, 1994.

Capítulo 26

La resolución de Montmort (1708, 1713) de los cinco problemas propuestos por Huygens en su tratado (1657)

Basulto Santos, Jesús
Pérez Hidalgo, Mª Dolores
Universidad de Sevilla

Introducción

Al final del tratado de Huygens, *De Ratiociniis in Ludo Aleae*, cuya versión en latín fue publicada por primera vez en 1657, encontramos cinco ejercicios propuestos por el autor y no resueltos, aunque en tres de ellos se da la solución.

Los cinco problemas se convirtieron en un desafío para los investigadores de aquella época. Autores como Hudde, Spinoza, Montmort, de Moivre, Jacques Bernoulli y Struyck se ocuparon de ellos de una manera u otra: resolviendo algunos de ellos o resolviéndolos todos. Sus soluciones, interpretaciones (con o sin reemplazamiento) y generalizaciones ejercieron una influencia incontestable en el desarrollo de la nueva disciplina a finales del siglo XVII y principios del XVIII. Ahora bien, este honor lo debe compartir Huygens con Fermat, que fue quién propuso el primero y el tercero, a través de Carcavy, en la carta que éste envió a Huygens el 22 de junio de 1656, y con Pascal, autor del quinto ejercicio propuesto por éste a Fermat, y que Huygens también conoció a través de la carta que Carcavy le envió el 28 de septiembre de ese mismo año (una traducción al español de la correspondencia indirecta entre Pascal y Fermat, con intermediación de Carcavy y con la participación de Huygens, durante los años 1655 y 1656, la encontramos en Basulto Santos, J., Camúñez Ruiz, J. A., 2007).

El propio Huygens, en años posteriores a la publicación de su tratado, fue resolviendo casi todos estos problemas (no se ha encontrado su resolución del tercero, si bien la solución del mismo la conocía el autor como se demuestra en la carta que éste envió a Carcavy el 6 de julio de 1656), a pesar de que sus resoluciones no vieron la luz hasta que no fueron recogidas en las Obras Completas del autor que la Sociedad Holandesa de las Ciencias publicó entre finales del siglo XIX y principios del XX.

Como se ha dicho, uno de los autores que abordaron la resolución de estos problemas fue Montmort en su importante tratado *Essai d'Analyse sur les Jeux de Hazard* (ediciones de 1708 y 1713). Según lo comentado más arriba, Montmort no conocía las resoluciones de Huygens de estos problemas, por lo que sus propias resoluciones pueden considerarse originales y, al disponer en la actualidad de lo escrito por ambos autores, tenemos la posibilidad de analizarlas y compararlas. Ese es el objetivo de este trabajo en el que, problema a problema, recogemos los enunciados bajo los que fueron presentados (curiosamente, hay algunos matices que diferencian los enunciados dados por ambos autores) y las reflexiones y cálculos que cada uno realizó para encontrar sus respectivas soluciones. Todo ello, bajo un lenguaje y notación más actual que la que nuestros autores usaron a mediados del XVII y principios del XVIII.

Problema primero

El primero de estos ejercicios es una generalización de la Proposición última del tratado de Huygens (Proposición XIV), la cuál sí estaba resuelta en el propio tratado. El enunciado bajo el que aparece en dicho tratado es:

> *A y B juegan juntos con dos dados con la condición siguiente: A habrá ganado si lanza 6 puntos, B si lanza 7. A hará en primer lugar un solo lanzamiento; a continuación B 2 lanzamientos sucesivos; después de nuevo A 2 lanzamientos; y así sucesivamente, hasta que uno u otro haya ganado. Se pide la relación de la chance de A a la de B. Respuesta: como 10355 es a 12276.*

En la segunda edición del libro de Montmort, la de 1713, éste aparece como Problema I o Proposición XXXII. El autor lo enuncia así:

> *Pedro y Pablo juegan juntos con dos dados: He aquí las condiciones del juego. Pedro ganará consiguiendo seis, y Pablo consiguiendo siete. Cada uno de ellos jugará dos lanzamientos sucesivos cuando tenga los dos dados: no obstante, Pedro, que comenzará, jugará con uno por primera vez. Se trata de determinar la suerte de cada uno de estos dos Jugadores, o la esperanza que cada uno tendrá de ganar la partida.*

La resolución aportada por Huygens aparece en la carta que envió a Carcavy el 6 de julio de 1656 y es muy parecida a la que da en la Proposición 14. Plantea la periodicidad del juego respecto de los jugadores que participan. Escribe que la misma seguirá la sucesión: ABBA ABBA … Esa periodicidad (y la independencia de los lanzamientos) le lleva a presentar una recurrencia bastante simple.

Tras enumerar los casos favorables a la obtención del "seis" al lanzar dos dados, y los favorables a la obtención del "siete", 5 y 6, respectivamente, llama x al valor del juego para A, o sea, la esperanza de este jugador al inicio del mismo y afirma que, en la situación de haberse realizado ya varios lanzamientos y en el momento en que A ha lanzado sin éxito el primero de los dos lanzamientos de su turno, entonces, antes de efectuar el segundo "*tendrá la misma apariencia de ganar que al inicio de juego*", o sea, x. Aquí es donde constatamos que Huygens se da cuenta de la periodicidad de este juego (ya había ocurrido en su resolución de la Proposición 14 del tratado). Eso le permite resolver el problema usando solamente las cuatro esperanzas e_1, e_2, e_3 y e_4, que son las valoraciones del juego para el primer jugador en las situaciones respectivas siguientas: al inicio del mismo, cuando se ha realizado la primera partida (cuando A ha lanzado sin éxito y es el turno de B), cuando se ha realizado la segunda partida (cuando B ha lanzado la primera de sus dos veces sin éxito), y cuando se ha realizado la tercera partida (cuando B ha realizado su segundo lanzamiento sin éxito y es ahora el turno de A). Llamando *a* al total apostado, esas cuatro esperanzas son:

- $e_1 = \frac{5}{36}a + \frac{31}{36}e_2$ (antes de lanzar, el jugador A tiene 5 posibilidades de ganar y 31 de que pase a ser el primer lanzamiento de B).
- $e_2 = \frac{6}{30}\cdot 0 + \frac{30}{36}\cdot e_3 = \frac{30}{36}e_3$ (en el primer lanzamiento de B, el jugador A tiene 6 posibilidades de perder, y 30 de que el jugador B disponga de su segundo lanzamiento).
- $e_3 = \frac{6}{36}\cdot 0 + \frac{30}{36}\cdot e_4 = \frac{30}{36}e_4$ (en el segundo lanzamiento de B, el jugador A tiene 6 posibilidades de perder y 30 de que se encuentre en el primero de sus dos lanzamientos).
- $e_4 = \frac{5}{36}a + \frac{31}{36}e_1$ (en el primero de los dos lanzamientos de A, este jugador tiene 5 posibilidades de ganar y 31 de disponer de un segundo lanzamiento y, por tanto, encontrarse como en el inicio del juego).

Resolviendo para e_1, que es realmente lo que nos interesa (esperanza del jugador A antes de iniciarse el juego), obtenemos $e_1 = \frac{10355}{22631}a$, por lo que la esperanza del segundo jugador en ese instante es la complementaria, o sea, $\frac{12276}{22631}a$. Por tanto, las suertes de ambos jugadores están en la relación 10355:12276, siendo la solución dada por Huygens.

Para este problema, Montmort ofrece una resolución casi idéntica. Está claro, en este caso, que aunque Montmort no conocía la de Huygens, si conocía la solución de su Proposición 14. Usa la expresión “método analítico” a esta forma de resolución, expresión que había sido acuñada ya por Huygens en su tratado. La novedad de Montmort en este problema es la nota que añade al final del mismo, en donde presenta la siguiente generalización, de la que da la solución sin explicación alguna:

> *El método analítico que se ha empleado aquí es siempre el mejor y el más corto, cuando ocurre que al cabo de un cierto número de lanzamientos los Jugadores se encuentran en el mismo estado en el que ellos estaban anteriormente; pero cuando esto no ocurre, se cae en series infinitas; y para encontrarlas, toda la habilidad consiste en observar bien las condiciones del juego, y en sacar la Ley de la progresión. Por ejemplo, si se supone que Pedro juega en primer lugar un lanzamiento, y Pablo dos lanzamientos, a continuación Pedro dos lanzamientos, y Pablo tres lanzamientos; a continuación Pedro tres lanzamientos, y Pablo cuatro lanzamientos, y así sucesivamente, Pablo jugando siempre un lanzamiento más que Pablo, la suerte de Pedro sería expresada por una serie en la que sería muy difícil obtener la suma, esta serie sería* $= \frac{b}{f}A + \frac{d\times c^2\times b}{f^4}A + \frac{d^2\times c^2\times b}{f^5}A + \frac{d^3\times c^3\times b}{f^9}A + \frac{d^4\times c^3\times b}{f^{10}}A + \frac{d^3\times c^5\times b}{f^{12}}A + \frac{d^8\times c^9\times b}{f^{16}}A + \frac{d^7\times c^9\times b}{f^{17}}A + \frac{d^8\times c^9\times b}{f^{18}}A + \frac{d^9\times c^9\times b}{f^{19}}A + \frac{d^{10}\times c^{14}\times b}{f^{25}}A +$ *& así, suponiendo* $b=5$, $c=30$, $d=31$, $f=36$.

> *Es fácil señalar el orden de la serie, y de continuar hasta el infinito, la expresión de la suerte de Pablo sería la cantidad que le falta a la serie que expresa la suerte de Pedro para valer A.*
> *Si Pedro y Pablo juegan con un dado, según el orden que se acaba de señalar, a que el primero consiga un seis, se tendrá para la expresión de la suerte de Pedro una serie más simple, a saber* $1-p+p^3-p^5+p^8-p^{11}+p^{15}-p^{19}+p^{24}-p^{29}+\&$ *así, suponiendo que sea* $p=\frac{1}{6}$.

Problema segundo

Este problema de Huygens tiene la originalidad de introducir por primera vez en la historia de esta disciplina, la distinción entre muestreo con y sin reemplazamiento. Eso ocurre en la discusión que surge entre este autor y su amigo Hudde sobre la resolución de este problema. El problema tuvo el siguiente enunciado para Huygens:

> *Tres jugadores A, B y C toman 12 fichas de las que 4 son blancas y 8 negras; juegan con esta condición de que ganará el que primero haya, escogiendo a ciegas, sacado una ficha blanca, y que A elegirá el primero, B a continuación, después C, después de nuevo A, y así sucesivamente, por turnos. Se pide la relación entre sus chances.*

Para Montmort, que lo presentó en su libro como Problema II, o Proposición XXXIII, el mismo mereció el siguiente enunciado:

> *Tres Jugadores, Pedro, Pablo y Jacobo juegan juntos y acuerdan que, sacando uno detrás del otro una ficha al azar entre doce, de las que ocho serán negras y cuatro blancas, el que primero extraiga una ficha blanca ganará. He aquí el orden según el cual ellos juegan: Pedro saca el primero, Pablo saca el segundo, y Jacobo el tercero; a continuación, Pedro vuelve a comenzar, y los otros le siguen según su orden, hasta que uno de los Jugadores haya ganado. Se trata de encontrar lo que cada Jugador debe poner en el juego, con el fin de que la partida sea igual.*

No es conocido de dónde derivó Huygens este segundo problema. Es probable que fuese inventado por él mismo. Jacques Bernoulli señala en la edición comentada del Tratado de Huygens (en la *Pars Prima* de su *Ars Conjectandi*) que hay tres posibles interpretaciones del problema: (i) cada bola o ficha negra es reemplazada después de la extracción, (ii) las bolas o fichas no son reemplazadas y hay una caja común que incluye inicialmente las doce y (iii) cada jugador tiene su propia caja con las doce bolas o fichas y van extrayendo, cada uno de su caja, por turno y sin reposición. Resulta que Huygens tenía la primera interpretación en mente, como se puede comprobar por la resolución que aparece en el Apéndice II del tratado. La Academia Holandesa de las Ciencias añadió al tratado de Huygens nueve apéndices en su edición del siglo XIX. En ellos se recogen las notas en cuadernos y hojas sueltas, que Huygens fue escribiendo a lo largo de su vida en relación al cálculo en juegos de azar. Esos cálculos y reflexiones sólo vieron la luz cuando fueron publicados en el siglo XIX, formando parte del Tomo XIV de las Obras Completas, por lo que Montmort los desconocía en el momento en que publica las dos ediciones de su tratado.

En su resolución, Huygens supone que las extracciones son con reemplazamiento, dado que a cada uno de ellos asigna cuatro posibilidades en una dirección y ocho en la contraria, o sea, la composición de la caja no se altera tras cada extracción. El orden de las extracciones es ABC ABC... Llama x, y, z a las esperanzas u oportunidades de los jugadores antes de iniciarse el juego (son los valores que se quieren conocer), y llama a a la apuesta. En la primera extracción, si se extrae blanca, el jugador A ganará el juego y si no, el jugador A se convierte en el tercero de la siguiente serie de extracciones, por lo que su oportunidad se convierte en z (esperanza del tercero en el turno de extracciones). Por tanto, A tiene 4 posibilidades de conseguir la apuesta a y 8 de conseguir z, por lo que, según la Proposición 3 de su Tratado (proposición clave en el desarrollo de todo el tratado), el valor del juego para A, al inicio del mismo, es $x = \frac{4a + 8z}{12}$.

Consideramos ahora la situación del jugador B también en el momento en que el juego se inicia. Este jugador tiene 4 posibilidades de conseguir 0 (si el jugador A acierta en la primera extracción) y 8 oportunidades de convertirse en el primer jugador de la nueva serie y, por tanto, tener de suerte x. Entonces, $y = \frac{0 + 8x}{12}$. Por último, la valoración del juego, al inicio del mismo, para el tercer jugador es: 4 posibilidades de conseguir 0 y 8 de convertirse en el segundo jugador de la serie de extracciones y, por tanto, tener derecho a y. Entonces, $z = \frac{0 + 8y}{12}$. Resolviendo: $x = \frac{9}{19}a,\quad y = \frac{6}{19}a,\quad z = \frac{4}{19}a$, o sea, las ratios de sus suertes son 9:6:4.

En general, si p es la probabilidad de éxito en una extracción (probabilidad que no varía por ser con reemplazamiento) y si $q = 1 - p$, se obtiene:

$$x = \frac{p}{1-q^3}a,\quad y = \frac{pq}{1-q^3}a,\quad z = \frac{pq^2}{1-q^3}a,$$

y las ratios de las suertes son $1 : q : q^2$ (Hald, 1990).

Montmort, como primera resolución, interpreta el problema en el mismo sentido que Huygens y resuelve de forma parecida, pero no igual. Se fija en el tercer jugador, estableciendo y construyendo la suerte del mismo en estas tres circunstancias: Cuando le toca al primero extraer, S, cuando le toca al segundo extraer, y, y su suerte cuando a él mismo le toca extraer, z. Entre estas suertes hay las siguientes relaciones (llama A al total del dinero que hay en juego):

- $S = \frac{4}{12} \cdot 0 + \frac{8}{12} \cdot y = \frac{2}{3} y$ (si gana el primer jugador, la suerte del tercero será cero, y si no gana, la suerte del tercero se convierte en la que le corresponde cuando el segundo va a extraer).
- $y = \frac{4}{12} \cdot 0 + \frac{8}{12} \cdot z = \frac{2}{3} z$ (si gana el segundo jugador en su extracción, la suerte del tercero se hace cero, y si no gana, la suerte del tercero se convierte en la que le corresponde cuando a él le toca extraer).
- $z = \frac{4}{12} \cdot A + \frac{8}{12} \cdot S = \frac{1}{3} A + \frac{2}{3} S$, (o sea, cuando es el turno de extraer del tercero tiene 4 posibilidades de ganar y, por tanto llevarse el dinero que está en juego, y 8 oportunidades de no ganar y, por tanto, pasar de nuevo a la situación en que le toca extraer al primero, en la que la suerte del tercero era S).

Resolviendo en las tres igualdades encuentra $S = \frac{4}{19}A$, lo que expresa la suerte del tercer jugador al inicio del juego, o sea, cuando va a extraer el primero.

A continuación procede de igual forma con el segundo jugador, llamando, *u* a la suerte de este jugador cuando al primero le toca extraer, obteniendo $u = \frac{9}{19}A$. Por último, obtiene la suerte del primer jugador restando a la cantidad total en juego, la suerte de los otros dos jugadores, o sea, realizando $A - \frac{4}{19}A - \frac{9}{19}A = \frac{6}{19}A$. Y concluye:

Por consiguiente, si se quiere que el juego sea de diecinueve escudos, será necesario que Pedro ponga nueve, Pablo seis, y Jacobo cuatro.

A continuación, este autor añade una nota en la que da la solución bajo la segunda interpretación de Bernoulli. En este caso no se detiene a explicar su procedimiento de resolución, aunque si se procede como en el caso anterior, construyendo la esperanza de cada jugador en cada una de las sucesivas extracciones, y teniendo en cuenta que en este caso, al ser sin reemplazamiento, el proceso tiene fin, se obtiene los resultados dados por el autor. Así es como escribe:

Si el sentido del problema es que cada Jugador después de haber sacado una ficha no la vuelve a poner más, se encontrará de la misma manera la suerte de los tres Jugadores, como estos tres números, 77, 53, 35.

Problema tercero

Como ya se ha dicho, este problema fue otro de los propuestos por Fermat a través de Carcavy. Su enunciado es:

A apuesta contra B, que de 40 cartas, donde hay diez de cada color, extraerá 4 de manera que tenga una de cada color. Se encuentra en este caso que la chance de A es a la de B como 1000 es a 8139.

La solución de Huygens, sin explicación alguna, aparece en la carta que éste envió a Carcavy el 6 de julio de 1656. En los apéndices y cartas publicados en sus obras completas, no se ha encontrado la solución de Huygens de este problema.

Montmort intercambió el orden de los problemas 3 y 4 de Huygens. Decidió mantener seguidos los problemas 2 y 4 por tratarse de una temática parecida. El enunciado con el que Montmort presenta este problema, bajo el título de Problema IV o Proposición XXXV es el siguiente:

Pedro apuesta contra Pablo que sacando, con los ojos cerrados, cuatro cartas entre cuarenta, a saber, diez rombos, diez corazones, diez picas y diez tréboles, sacará una de cada clase. Se pide cuál es la suerte de estos dos Jugadores, o lo que ellos deben poner en el juego para apostar con igualdad.

En este problema, Montmort rompe su línea de resolución que, en los anteriores, había sido similar a la de Huygens (lo que este último llamó método analítico). En este caso, Montmort hace uso de un método que encontraba más eficaz y rápido: usando la combinatoria. Comienza por recordar una proposición que había demostrado al principio de su tratado (Proposición VII) cuyo enunciado es el siguiente:

PROPOSICIÓN VII.

Pedro, teniendo entre sus manos un número cualquiera de fichas de todos los colores, blancas, negras, rojas, verdes, & así, apuesta contra Pablo, que sacando al azar un número cualquiera determinado de fichas, él sacará tantas blancas, tantas negras, tantas rojas, tantas verdes, & así. Se pide cuántos azares tiene Pedro para hacer lo que se propone.

SOLUCIÓN.

Es necesario multiplicar el número que expresa de cuántas maneras las fichas blancas que Pedro debe tomar al azar, pueden ser tomadas de forma diferente en el número de fichas blancas propuestas, por el número que expresa de cuántas maneras las fichas negras que Pedro debe tomar al azar, pueden ser tomadas de forma diferente en el número entero de fichas negras propuestas; multiplicar a continuación este producto por el número que expresa de cuántas maneras diferentes las fichas rojas que Pedro se propone extraer, pueden ser tomadas en las fichas rojas propuestas, multiplicar de nuevo ese producto por el número que expresa de cuántas maneras diferentes las fichas verdes que se piden pueden ser tomadas entre todas las verdes, & así, sucesivamente, se tendrá el número buscado.

Esta solución lleva consigo la demostración, y no tiene dificultad alguna; pero como este Teorema será un resultado de un gran uso.

Entonces, Montmort comenta que este problema *no es más que un caso particular de la Proposición VII* y, por lo tanto, la suerte de Pedro es

$$\frac{\binom{10}{1}\cdot\binom{10}{1}\cdot\binom{10}{1}\cdot\binom{10}{1}}{\binom{40}{4}}=\frac{10^4}{\binom{40}{4}}=\frac{10000}{91390},$$

o sea, las suertes de ambos jugadores están en la relación de 10000 a 81390, con lo que propone una solución muy actual, tal y como hoy día se suele resolver este problema. Entonces añade dos ejemplos más con sus soluciones:

Si se pidiese cuánto hay que apostar que Pablo extrayendo trece cartas al azar en cincuenta y dos, no sacase todas de un color, se encontraría que hay que apostar 158753389899 contra 1.

Si se quiere saber cuánto hay que apostar que Pedro, extrayendo diez cartas al azar entre cuarenta, a saber, un as, un dos, un tres, un cuatro, un cinco, un seis, un siete, un ocho, un nueve y un diez de rombos, tantas de corazones, de picas y de tréboles, él sacará una decena completa, se encontrará que hay que apostar 1048576 contra 846611952, más o menos, como 1 contra 808.

Problema cuarto

Como ya se ha dicho, este problema tiene un contexto similar al segundo y también admite una doble interpretación, entendiendo nuestro autor las extracciones sin reemplazamiento. Su enunciado es:

Se toma, como más arriba, 12 fichas de las que 4 son blancas y 8 negras. A apuesta contra B que entre 7 fichas que él sacará a ciegas, se encontrarán 3. Se pide la relación entre la chance de A y la de B.

La solución de Huygens aparece también en el Apéndice II antes citado. En el mismo, Huygens considera también la modificación propuesta por Hudde en la que el resultado puede ser tres o más fichas blancas. En el segundo apartado de este apéndice Huygens lo enuncia de la siguiente forma:

A apuesta contra B que entre 12 fichas, de las cuales 4 son blancas y 8 negras, él tomará a ciegas 7 fichas, de las cuales 3 serán blancas, y no más. Se pide la relación de la chance de A a la de B. Respuesta: como 35 es a 64.

Vemos que el enunciado es idéntico al cuarto problema. Solamente, a consecuencia de un malentendido que había tenido lugar entre él y Hudde, sobre la interpretación del enunciado, Huygens añadió al mismo las palabras “en niet meer”, que indican que para ganar A debe tener 3 fichas blancas “y no más”, dado que Hudde entendía el enunciado como “extraer al menos 3 blancas”. Huygens supone que las fichas son extraídas una tras otra y comienza por calcular la esperanza o suerte de A, después de la sexta extracción, en los dos únicos casos donde éste puede ganar en la séptima. A continuación considera la situación del juego después de la quinta extracción, y así sucesivamente, para remontarse por fin al inicio del juego. La resolución puede contemplarse de la siguiente forma: En la situación inicial hay 4 blancas y 8 negras. Supongamos el instante en el que se han extraído b blancas y n negras, quedando entonces en la caja $4-b$ blancas y $8-n$ negras (en total quedan $12-b-n$ fichas). Sea $e(b,n)$ el valor de la suerte o esperanza del primer jugador en ese instante. En la siguiente extracción puede salir una ficha blanca (hay $4-b$ posibilidades de que eso ocurra) y pasar así a un juego cuya valoración para el jugador A es $e(b+1,n)$, o puede salir negra (hay $8-n$ oportunidades para ello) y encontrarse entonces en un juego cuya valoración para el primer jugador es $e(b,n+1)$. Aplicando la Proposición 3 del tratado de Huygens:

$$e(b,n)=\frac{(4-b)\cdot e(b+1,n)+(8-n)\cdot e(b,n+1)}{12-b-n}$$, donde $0\le b\le 4$ y $0\le b+n\le 7$.

Desde luego, si el jugador A consigue su objetivo (3 blancas y 4 negras) gana la apuesta, o sea, $e(3,4)=a$. Ahora bien, $e(b,7-b)=0$, en cualquier otro caso, pues culminadas las 7 extracciones, el primer jugador sólo gana cuando consigue la extracción (3,4). En la situación “3 blancas y 3 negras” (ha extraído 6 fichas y le falta una por sacar que si es blanca gana), Huygens obtiene para el jugador A: $e(3,3)=\frac{5}{6}a$.

Pues bien, partiendo de esta igualdad y analizando todas las posibilidades favorables a ese jugador, hacia atrás y tras 19 igualdades, Huygens llega a la situación (0,0), o sea, el punto de partida, para el que encuentra $e(0,0)=\frac{35}{99}a$, por lo que concluye que la chance de A es a la de B como 35 es a 64. Es claro, en este caso, que el método analítico Huygens se hace largo y tedioso.

Por último, en el mismo apéndice, Huygens aborda este problema otra vez pero introduciendo en el enunciado la modificación propuesta por Hudde. O sea, el jugador A debe conseguir 3 o más fichas blancas en las 7 extracciones. Por tanto, los resultados que le hacen ganar el juego son (3,4) y (4,3).

Teniendo el precedente del apartado anterior, Huygens simplifica el cálculo reemplazando el problema en cuestión por el "problema complementario". Según éste, el jugador A debe tomar 5 fichas de las que cuatro al menos deben ser negras. Señalemos que no es necesario discutir el caso de 4 fichas de las que 0 son blancas y 4 negras, porque entonces la quinta extracción siempre hará ganar al jugador A. En este caso, tras ocho igualdades, llega a la solución 42 a 57, o 14 a 9.

Este problema es presentado por Montmort en su texto como Problema III o Proposición XXXIV. El enunciado y resolución van como sigue:

> *Pedro apuesta contra Pablo que cogiendo, con los ojos cerrados, siete fichas entre doce, de las que ocho son negras y cuatro blancas, él cogerá tres blancas y cuatro negras. Se pide cuánto deben apostar Pedro y Pablo para que la apuesta de cada uno esté en la misma proporción que su suerte.*

En este caso, Montmort hace uso de nuevo de su Proposición VII, o sea de la combinatoria, y resuelve el problema de una forma sencilla y actual. La suerte del primer jugador es calculada mediante $\dfrac{\binom{4}{3}\binom{8}{4}}{\binom{12}{7}} = \dfrac{35}{99}$, por lo que la del segundo jugador es $\dfrac{64}{99}$.

Y el autor añade la solución para el caso de tres o más fichas blancas: *Si se quiere que Pedro haya ganado también cuando coja cuatro blancas y tres negras, se tendrá de igual forma, por el art. 20,* $\dfrac{70\times 4+1\times 56}{792}=\dfrac{14}{33}$ *para la suerte de Pedro, y en este caso sería necesario que Pablo ponga en el juego 19 contra 14 de Pedro.* O sea, en este caso calcula

$$\frac{\binom{4}{3}\binom{8}{4}}{\binom{12}{7}}+\frac{\binom{4}{4}\binom{8}{3}}{\binom{12}{7}}.$$

Problema quinto

Éste es el problema propuesto por Pascal ("que juzga más difícil que todos los demás") a Fermat y, a través de Carcavy, a Huygens, en una carta de 28 de septiembre de 1656. Esta carta contenía la solución sin explicación de los dos sabios franceses. La solución de Huygens aparece en su respuesta a Carcavy de 12 de octubre de 1656. El problema es conocido como el de la Ruina del Jugador y el enunciado que aparece al final del Tratado de Huygens es el siguiente:

> *Habiendo tomado cada uno 12 fichas, A y B juegan con 3 dados con esta condición de que a cada tirada de 11 puntos, A debe dar una ficha a B, y que B debe dar una ficha a A en cada tirada de 14 puntos, y que ganará aquel que sea el primero en poseer todas las fichas. Se encuentra en este caso que la chance de A es a la de B como 244140625 es 282429536481.*
>
> La resolución detallada del problema aparece en unas hojas sueltas escritas por Huygens en 1676 y que la edición de sus obras la incluye como Apéndice VI. Queremos señalar que estas hojas sueltas son presentadas como las primeras en la historia que incluyen árboles de probabilidad para la comprensión y demostración de los resultados.

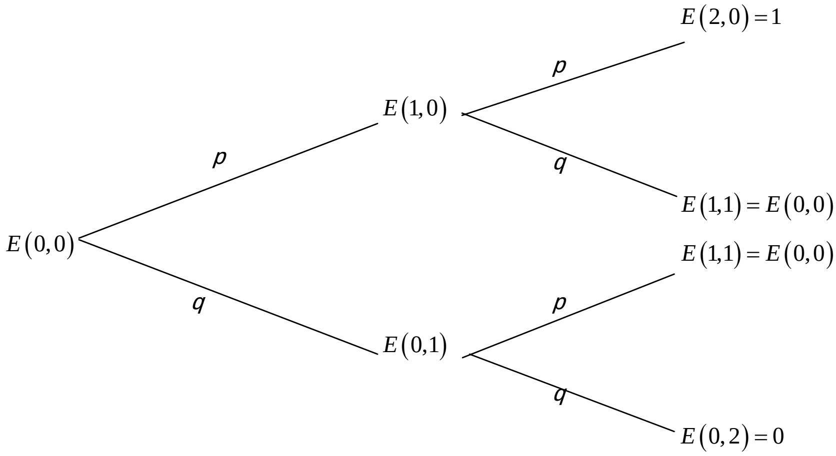
El enunciado tal y como fue presentado inicialmente sería: Sean dos hombres que juegan con tres dados, donde el primer jugador consigue un punto si se lanza 11 y el segundo lo consigue si se lanza 14. Pero en lugar de que los puntos se acumulen en la vía ordinaria, dejamos que un punto sea añadido al tanteo de un jugador sólo si el tanteo de su oponente es cero y, en otro caso, dejamos en su lugar que el punto sea sustraído del tanteo de su oponente. Es como si los puntos opuestos formasen "pareja" y se aniquilasen unos a otros, por lo que el jugador que va detrás siempre tiene cero puntos. El ganador es el primero que alcance 12 puntos. ¿Cuál es la chance de cada jugador?

Esta es una versión libre de la descripción dada por Carcavy y no hay razón para suponerla muy diferente de la que Pascal había propuesto. Cuando Huygens la lee inmediatamente la piensa en términos de que los puntos de los jugadores se acumulan en la vía ordinaria, pero el ganador será el primero que lleve doce puntos de ventaja (en 1656), y cuando él lo plantea en su De ratiociniis in aleae ludo *(en 1657) lo da en términos de que cada jugador arranca con doce puntos, y una ganancia de un jugador supone la transferencia de un punto de su oponente a él mismo, y el ganador total es el que arruina al otro de todos sus puntos. Las tres formas de este problema de Pascal son, desde luego, equivalentes, pero es en la última forma como el problema es conocido como el de la "Ruina del Jugador".*

Entonces, en la formulación de Pascal, los jugadores arrancan con el tanteo (0,0), y el ganador es aquel que primero consigue 12 puntos (teniendo el otro 0 puntos). Pues bien, ésta es también la forma que Huygens usa en su resolución de 1676.

El número de chances para ganar un punto es 15 para A y 27 para B; tomaremos $p = 15/42 = 5/14$ y $q = 27/42 = 9/14$. Supongamos que $E(a,b)$ es la chance que tiene A de ganar cuando este jugador tiene a puntos y B tiene b puntos. El problema es encontrar $E(0,0)$.

Huygens comienza analizando el caso simple en el que el juego se acaba cuando uno de los jugadores llega a 2 puntos. Da una lista de todos los posibles resultados y sus probabilidades en un diagrama de tipo árbol. El diagrama que presenta, con un lenguaje actual, es el siguiente:



Aquí, el total apostado es tomado como 1. Del esquema podemos deducir las igualdades:

$$E(0,0)=p\cdot E(1,0)+q\cdot E(0,1)==p\big(p\cdot E(2,0)+q\cdot E(1,1)\big)+q\big(p\cdot E(1,1)+q\cdot E(0,2)\big)=$$

$$=p\big(p+q\cdot E(0,0)\big)+q\big(p\cdot E(0,0)+q\cdot 0\big)==p^2+2pq\cdot E(0,0).$$

Teniendo en cuenta que $p+q=1$, podemos despejar $E(0,0)$, obteniendo: $E(0,0)=\dfrac{p^2}{p^2+q^2}$. Por tanto, la esperanza del primero es a la del segundo como p^2 es a q^2.

Después, Huygens estudia el caso en el que el ganador ha de conseguir cuatro puntos de ventaja. Ingeniosamente resuelve considerando sólo uno de cada dos posibles estados del juego, a saber los puntos (4,0), (2,0), (0,0), (0,2) y (0,4), y señala que el árbol de sucesos será similar al primero, salvo que todas las chances son cuadradas. La justificación para omitir los puntos intermedios es, evidentemente, que para pasar de (0,0) a (0,4), por ejemplo, es necesario pasar por (0,2). Huygens no ofrece explicación, más allá de un diagrama, aunque el planteamiento origina tres ecuaciones con la solución $E(0,0)=\dfrac{p^4}{p^4+q^4}$. Entonces, las chances de los dos jugadores están en la relación $p^4:q^4$.

Finalmente, señala que si es necesaria una ventaja de 8 puntos, aplicando nuevamente el argumento anterior, $E(0,0)=\dfrac{p^8}{p^8+q^8}$, y así.

Si se requiere una ventaja de 3 puntos para ganar, él hace el paso de (0,0) a (1,0) con probabilidad p y después a (3,0) con probabilidad p^2, y de manera similar para las demás ramas del diagrama, lo que le lleva a las ecuaciones

$$E(1,0)=\frac{p^2E(3,0)+q^2E(1,2)}{p^2+q^2}=\frac{p^2+q^2E(0,1)}{p^2+q^2},$$

$$E(0,1)=\frac{p^2E(2,1)+q^2E(0,3)}{p^2+q^2}=\frac{p^2E(1,0)}{p^2+q^2},$$

$$E(0,0)=p\cdot E(1,0)+q\cdot E(0,1),$$

con la solución $E(0,0)=\dfrac{p^3}{p^3+q^3}$. De aquí, él generaliza para $\dfrac{p^6}{p^6+q^6}$.

Los casos considerados hasta aquí requieren la solución de tres ecuaciones. Si es necesaria una ventaja de 5 puntos para ganar, el número de ecuaciones llega a ser considerablemente mayor, y Huygens señala que una solución puede ser obtenida como para $n=3$ pero que tomaría un tiempo mucho mayor. Finalmente, afirma que, en general, la ratio de las esperanzas de A a B es $p^n:q^n$, aunque *"no vemos como concluir en general que las esperanzas de A y B están en la razón de las potencias"*. La respuesta dada al final del problema 5 será entonces $5^{12}:9^{12}$.

El enunciado y resolución de Montmort es el siguiente:

Pedro y Pablo cogen cada uno doce fichas y juegan con tres dados con las condiciones que siguen. Si los dados llevan once, Pablo dará una ficha a Pedro. Si los dados llevan catorce, Pedro dará una ficha a Pablo. Aquél de los dos que primero tenga todas las fichas, ganará. Se pide cuál es la suerte de los dos Jugadores.

Para resolver, en primer lugar Montmort calcula la suerte de cada jugador en cualquier lanzamiento de los tres dados. El número de posibles resultados es $6^3=216$. Los favorables al primer jugador, que debe sumar once puntos entre los tres dados, son 27, y los que favorecen al segundo, que deben sumar catorce, son 15. El resto hasta 216, o sea, 174, no favorecen ni a uno ni a otro. Por tanto, Montmort maneja las fracciones $\dfrac{27}{216}$, $\dfrac{15}{216}$ y $\dfrac{174}{216}$ en la resolución de este problema.

A continuación, el autor define las "suertes" del primer jugador cuando él mantiene las 12 fichas y le ha ganado al contrario 1, o 2, o 3, …, o las 12 al segundo jugador, y al contrario, cuando el contrario mantiene las 12 y él va perdiendo 1, o 2, o 3,…, o las 12. Así, por ejemplo, si

x: suerte del primer jugador cuando él tiene 12 fichas y el segundo también 12,

y: suerte del primer jugador cuando él tiene 12 fichas y el segundo tiene 11,

k: suerte del primer jugador cunado él tiene 11 fichas y el segundo 12, se verifica:

$x=\dfrac{27}{216}y+\dfrac{15}{216}k+\dfrac{174}{216}x$, o sea, $14x=9y+5k$. De esta forma llega a plantear otras 19 igualdades que tras un proceso de sustitución hacia atrás le lleva a (siendo *A* el dinero total en juego): *Y como consecuencia se encontrará* $x=\dfrac{9y+5k}{14}=$

$\frac{282429536481A+352487604195x}{987648885496}$, *sustituyendo para y & k sus valores en x, y por fin reduciendo se tendrá* $x=\frac{282429536481}{282673677106}A$, *lo que expresa la suerte de Pedro, y* $A-x=\frac{244140625}{282673677106}$, *lo que expresa la suerte de Pablo.*

Montmort termina este problema añadiendo esta nota en la segunda edición de su texto:

Es a propósito observar que la vía analítica no es aquí, quizás, la mejor, puesto que se puede descubrir de otra forma que las suertes de Pedro y de Pablo son como las doceavas potencias de los números 9 y 5, así como ha sido observado por los señores Bernoulli que me han avisado en sus cartas del 17 de marzo de 1710 y 26 de febrero de 1711, y después por el señor de Moivre en su Tratado De Mensura Sortis, que apareció el año pasado.

Conclusión

Este autor, P. R. de Montmort, en su excelente tratado sobre el cálculo en juegos de azar, incorpora como parte del mismo los cinco problemas propuestos por Huygens casi cincuenta años antes, los resuelve y, como corresponde al proceso natural del crecimiento matemático, introduce generalizaciones de los mismos. En dos de estos problemas utiliza el método más natural y sencillo de resolución, la combinatoria, método actual por otra parte, y en los otros tres usa el de su predecesor, el "método analítico", reconociendo lo farragoso que llega a resultar, sobre todo, en la resolución del último problema.

Bibliografía

BASULTO SANTOS, J., CAMÚÑEZ RUIZ, J. A., ORTEGA IRIZO, F. J., PÉREZ HIDALGO, M. D. (2006). "El problema de los puntos para jugadores con desigual destreza: la solución de Montmort (1713)". Capítulo 2, pag. 13-22. *Historia de la Probabilidad y la Estadística (III)*. Delta Publicaciones. Madrid

BASULTO SANTOS, J., CAMÚÑEZ RUIZ, J. A., (2007). *La geometría del azar.* Nivola, libro y ediciones. Madrid.

CAMÚÑEZ RUIZ, J. A. (2004). *La Probabilidad y la Estadística en el Período 1654-1670. Sus Antecedentes en el Renacimiento Italiano*. Tesis Doctoral. Universidad de Sevilla.

DAVID, F. N. (1962) . "Games, Gods and Gambling".Charles Griffin & Co. Ltd., London.

EDWARDS, A.W.F. (1983) "Pascal's Problem: The "Gambler's Ruin" Int. Statist. Re. *51, 73-79.*

FREUDENTHAL, H. (1980). "Huygens' Foundations of Probability".*Historia Matemática*, 7, 113-117.

HALD, A. (1990). *A History of Probability and Statistics and Their Applications before 1750.* John Wiley & Sons. New York.

HUYGENS, C. *Oeuvres Complètes*. 22 volúmenes. Société Hollandaise des Sciences. Nijhoff, La Haye. 1888-1950. Los volúmenes usados aquí son: Vol. I y XIV.

KENDALL, M.G. (1956)."The beginnings of a probability calculus".*Biometrika,* 43, 1-14.

KENDALL, M.G., PLACKETT, R. L. (1977). *Studies in the History of Statistics and Probability. Vol II.* Griffin, London.

MONTMORT, P. R. DE (1713) *Essay d'Analyse sur les Jeux de Hazard*. Seconde Edition. Revûe et augmentée de plusieurs Lettre. Quillau, París. Publicado de forma anónima.

REIERSOL, O. (1968). "Notes on some propositions of Huygens in the Calculus of Probability". *Nordisk Matematsk Tidskrift*, 16, 88-91.

SCHNEIDER, I. (1980). "Christiaan Huygens's Contribution to the Development of a Calculus of Probabilities". *Janus* LXVII, 269-279.

TODHUNTER, I. (1865). *A History of the of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace.* Macmillan, London. Reprinted by Chelsea, New York, 1949.

Capítulo 27

Aportación al conocimiento de la obra de Manuel Escudé Bartolí (1856-1930). Impulsor de la estadística municipal en Barcelona

TERESA CORBELLA DOMÈNECH
Universidad Rovira Virgil
MANUEL ESCUDÉ AIXELÀ
Universidad de Barcelona

Introducción

Manuel Escudé Bartolí (Reus, 1856 – Barcelona, 1930) fue el primer director del entonces llamado Negociado de Estadística, Padrón y Elecciones del Ayuntamiento de Barcelona, órgano creado en 1902 que tenía que paliar la, hasta aquel entonces, falta de recogida sistemática y regular de estadísticas de una ciudad que estaba creciendo y contaba ya más de medio millón de habitantes. Bajo su dirección se publicaron 17 volúmenes de los Anuarios Estadísticos de la ciudad de Barcelona.

Sin embargo, aunque su trabajo de estadístico es de especial relevancia, los intereses y preocupaciones de este hombre se dirigen también hacia otros campos. Así, en su obra, que es extensa, desarrolla su vertiente de geógrafo y deja traslucir sus preocupaciones por temas sociales.

En este trabajo se presenta a Manuel Escudé Bartolí. Primero se anotan sus datos personales, luego se repasa su trayectoria profesional dividiéndola en dos partes: a) el periodo anterior a su incorporación al Ayuntamiento de Barcelona, b) la etapa en el Ayuntamiento de Barcelona como jefe del Negociado de Estadística, Padrón y Elecciones que perdura hasta su jubilación en 1923. A modo de conclusión se comentan las repercusiones del legado de Manuel Escudé Bartolí. En el quinto apartado se recogen sus escritos más importantes. Cabe señalar que además de monografías y anuarios también había escrito numerosos artículos de periódico.

Datos personales

Manuel Escudé Bartolí nace en Reus el 1 de octubre de 1856. Su padre, Antonio Escudé, era un carpintero reusense; su madre, Maria José Bartolí, era oriunda de un pueblo cercano: La Selva del Camp. La familia Escudé Bartolí vivía en la calle de Jesús. De su etapa inicial en Reus no se hallan muchos datos [pueden consultarse las reseñas bibliográficas de Gras (1899, p. 81) o Olesti (1991, p. 228)]. Sin embargo, cabe destacar que de joven colaboraba con la prensa local, así escribía en *El Eco del Centro de Lectura*[1]. De hecho, Olesti, en su diccionario de reusenses, le considera geógrafo y escritor. En cambio, en la reseña que se le dedica en la Enciclopedia Universal Ilustrada Espasa-Calpe aparece como publicista, y en la Gran Enciclopedia Catalana es un estadístico y demógrafo (Alier, 1974).

De la familia de Manuel Escudé Bartolí cabe mencionar a su hermano José (1863-1898) que desgraciadamente murió con solo 34 años, pero había iniciado una prometedora carrera como pintor (Ràfols, 1980). También debe mencionarse a su mujer Josefa Molist Carbó que era prima por lado materno de Benet Chías Carbó. Benet Chías Carbó colaboró profesionalmente en diversas ocasiones en su calidad de cartógrafo con Manuel Escudé Bartolí; así, en el "Atlas geográfico de España" (1901) publicación en fascículos que a partir del fascículo 63 pasó a llamarse "Atlas geográfico ibero-americano" y fue reeditado en 1904. Esta obra fue declarada de utilidad pública para la enseñanza el 2 de enero de 1904.

Al salir de Reus, se traslada a Madrid y gana una oposición para el cuerpo de estadística del Instituto Geográfico y Estadístico. Es destinado a la delegación de Barcelona (Gras, 1899). En aquel periodo colabora en el censo de 1877. En 1878, gana una plaza por oposición en el mismo cuerpo como auxiliar de segunda, el cuerpo de facultativos estadísticos era de recién creación[2]. Años más tarde acaba siendo el jefe de la delegación de estadística de Barcelona.

En 1902 se crea el Negociado de Estadística, Padrón y Elecciones en el Ayuntamiento de Barcelona. El Consistorio opta por poner una persona ajena al Ayuntamiento al frente, la persona elegida es Manuel Escudé Bartolí. En 1906 se le conceden los honores de jefe de la Administración Civil. Se jubila de su cargo en el Ayuntamiento en 1923. Muere en Barcelona el 29 de noviembre de 1930.

Trayectoria profesional

A- Del Instituto Geográfico y Estadístico al Ayuntamiento de Barcelona

Tras pasar su infancia en Reus donde cursa su bachillerato y colabora como articulista en la prensa local, se traslada a Madrid. Allí entra a formar parte de la primera generación de estadísticos que dependían del Instituto Geográfico y Estadístico, y colabora en la tabulación y la depuración de los datos del censo de 1877. Era el tercer censo que se realizaba en poco tiempo en el país (1857 y 1860), pero era el primero que contaba con un cuerpo específico de técnicos para su realización.

Poco después, en 1880 publica el extenso "Diccionario estadístico español de los resultados generales del censo de la población, según el empadronamiento hecho en 31 de diciembre de 1877 y declarados oficiales por R.O. de 18 de abril de 1879". No es su primera

---

[1] El Centro de Lectura era una de las grandes instituciones culturales de su ciudad natal, que todavía funciona hoy en día.
[2] El Instituto Geográfico y Estadístico se crea en 1873 (Arqüeso y Celestino, 1999).

monografía, dos años antes ya había salido el folleto de 25 páginas “Estudios sobre la historia física de la tierra”, y ese mismo año también aparece “Las maravillas del mar”. En aquella época también era corresponsal de los diarios de la Corte *El Día* y la *Nueva Prensa*.

En 1878 gana una oposición como auxiliar de segunda en el Instituto Geográfico y Estadístico. Posteriormente, en 1887 asciende a auxiliar de primera y a oficial de tercera en 1894. En 1897 se convierte en el jefe de la delegación provincial de Barcelona que se ocupa de los trabajos estadísticos de la zona. Olesti (1991) comenta que también había pasado por la delegación de Gerona.

A lo largo de toda esta etapa nunca abandonó su vertiente de geógrafo al que le gusta escribir. En 1885 publica “Las Carolinas. Descripción geográfica y estadística del archipiélago carolino. 1ª parte” que lleva un efusivo antetítulo “Nuestras colonias”. En 1891 sale de la imprenta un trabajo de unas 300 páginas en el que Escudé colabora con Luís de Ramón bajo la dirección de Pablo Riera “España y sus colonias: noticias de su población, agricultura, industria y comercio, según los más recientes datos estadísticos”. Seis años después, en 1897, otra monografía “La producción española en el siglo XIX: estadística razonada y comparada” de unas 300 páginas sale de una imprenta barcelonesa. Cabe señalar que las imprentas son, ya desde el principio, barcelonesas.

En 1891 había aparecido otro libreto de 32 páginas “España social y económica”, pero en este caso se trataba de una recopilación de artículos aparecidos en el *Diario Mercantil*. También es redactor jefe de la revista *La Exposición* y de *Revista Geográfica y Estadística* que es una de las primeras publicaciones periódicas aparecidas en Barcelona con una vertiente difusora de datos (Gras, 1899; Olesti, 1991). Cabe destacar también su participación en la revista *El Trabajo Nacional*, Urteaga y Nadal (2002) consideran que hace de director virtual. Esta revista, fundada en 1892, es el órgano de expresión de la organización patronal Fomento del Trabajo Nacional y en ella se defienden los intereses proteccionistas de la patronal. En agosto de 1902 Manuel Escudé Bartolí es nombrado vicesecretario de dicha organización (Urteaga y Nadal, 2002).

Su actividad investigadora también le lleva a participar en el Congreso Internacional de Higiene y Estadística celebrado en Madrid en 1898 donde es vicepresidente de una sección del congreso y presenta “Memoria estadística de la población de Barcelona”. A principios del siglo XX, antes de entrar en el Ayuntamiento de Barcelona presenta “La natalidad en Barcelona” (1901) en la que muestra su preocupación por la baja natalidad en la ciudad y permite ejemplificar su interés por temas de carácter social.

A finales del siglo XIX Manuel Escudé Bartolí era una personalidad consagrada y con contactos dentro de la estadística catalana. Su trayectoria profesional era sólida, su obra entendida como investigaciones estaba en pleno desarrollo y sus intereses iban mucho más allá de las cuestiones técnicas.

B- En el Ayuntamiento de Barcelona

i) El contexto municipal

Las elecciones municipales de noviembre de 1901 cambiaron el signo del Ayuntamiento de Barcelona. El nuevo consistorio desea modernizar la institución e impulsa una reestructuración de la administración municipal, por otro lado necesaria, dadas las nuevas necesidades.

En 1902 Barcelona cuenta con 546.982 habitantes[3], acaba de organizar la exposición de 1888 y está en plena expansión tanto en lo referente al término municipal como al número de habitantes. El número de habitantes en 1910, 1920 y 1930 (años en los que se efectúan censos) son respectivamente: 587.491, 710.335 y 1.005.565 habitantes. En cuanto al incremento del territorio cabe recordar las anexiones de los pueblos de Gràcia, Sants, Les Corts, Sant Martí de Provençals, Sant Gervasi y Sant Andreu por R.D. de 20 de abril de 1897 que cuadriplicaron el tamaño de la ciudad pasando de unos 15 km^2 a unos 60 km^2, la anexión de Horta por R.D. de 3 de Julio de 1903 y la de la actualmente llamada Zona Franca por ley de 11 de mayo de 1920 que incrementaron cada una de ellas el municipio en aproximadamente 10 km^2 y, finalmente, la última gran anexión[4] fue la de Sarria a 4 de noviembre de 1921 configurando una ciudad de 96 km^2. En veinte y pocos años la ciudad vio su extensión multiplicada por más de seis. La densidad de población por aquel entonces rondaba los 9.000 habitantes/km^2.

En este contexto cabe decir que las necesidades informativas del Consistorio eran importantes. Sin embargo, no se puede considerar que existiera realmente un servicio de estadística municipal. Había habido algunas estadísticas puntuales de iniciativa municipal en el campo de la estadística financiera en 1877 o de la estadística demográfica en 1882 (Ventura, 2002, p. 15). Tatjer (2001) comenta las iniciativas municipales para crear un registro civil que se abrió definitivamente en 1841. La única excepción a la elaboración continuada de estadísticas municipales fue en el campo sanitario. Desde 1888 el cuerpo de médicos municipales editaba la *Gaceta Sanitaria de Barcelona* que recogía información sobre el movimiento de población, el registro de enfermedades e información meteorológica. En 1891 se crea el Instituto de Higiene Urbana que entre otras funciones tiene la de encargarse de la estadística demográfico-sanitaria de la ciudad.

Sin embargo, el hecho que prácticamente no se hicieran estadísticas desde el Ayuntamiento no significa que no haya ningún estudio a lo largo del siglo XIX en el que no se recogieran y utilizaran estadísticas en la ciudad de Bacelona. El más conocido de estas estudios es seguramente el realizado por Laureano Figuerola (1849) “Estadísticas de Barcelona en 1849”. Otro trabajo de obligada mención es “Teoría general de la urbanización, y aplicación de sus principios y doctrinas a la reforma y ensanche de Barcelona” de Ildefonso Cerdà (1867) cuyo segundo volumen es “La urbanización considerada como un hecho concreto. Estadística urbana de Barcelona”. Ambos libros merecieron ser reimpresos en la década de los sesenta (siglo XX) por el Instituto de Estudios Fiscales. Mucho menos conocida, pero que por su carácter de estudio realizado en el marco de una compañía privada es interesante, es el trabajo de Arró (1892) “Estadística médica de la Compañía de los Ferrocarriles de Tarragona a Barcelona y Francia correspondiente al septenio de 1879 a 1885 respecto de las líneas de Gerona y al año 1886 respecto de éstas y de las de Tarragona y la Frontera”, también reimpreso en 1985 por la Universidad de Barcelona.

ii) El Negociado de Estadística, Padrón y Elecciones

En enero de 1902 se constituye el nuevo consistorio. Este crea dos nuevos negociados el de Ingresos y el de Estadística, Padrón y Elecciones (posteriormente llamado Negociado de Estadística). El primero debía sanear el estado de cuentas municipales, el segundo debía proveer el primero de información y además elaborar unas estadísticas dignas de una ciudad grande.

[3] La fuente de todos los datos sobre la ciudad es el Ayuntamiento de Barcelona.
[4] Ha habido alguna pequeña modificación posterior pero la extensión del término municipal no ha cambiado significativamente.

Explícitamente, se buscó para la dirección de estos negociados personas ajenas al Ayuntamiento que tuvieran demostrada experiencia en sus campos. El concurso para elegir al jefe del Negociado de Estadística se resolvió el 25 de octubre de 1902 a favor de Manuel Escudé Bartolí quien tomo posesión del cargo el 1 de diciembre siguiente y lo conservó hasta el día de su jubilación el 31 de octubre de 1923 (con cambio de nombre y alguna reorganización). El Negociado de Estadística se ocupó de las estadísticas municipales con excepción de las estadísticas sanitarias de las que se siguió encargando el Instituto de Higiene Urbana.

Bajo la dirección de Manuel Escudé Bartolí ya desde 1902 se publicaron los “Anuarios Estadísticos de la ciudad de Barcelona”. Así aparecieron un total de 17 volúmenes, uno por año desde 1902 hasta 1917, y uno final que recogía el trienio 1918-1920 editado en 1923. Se trata de unos volúmenes de gran formato, bien impresos y de buena presencia que podían presentarse a otras instituciones. Están divididos en varios apartados, la estructura inicial se acordó en la Comisión de Estadística consistorial: territorio, población, natalidad, nupcialidad, mortalidad, vigilancia y seguridad, policía urbana, instrucción pública, museos municipales y concursos de edificios, beneficencia, justicia, cementerios, riqueza imponible, hacienda municipal, abastos, industrias, comercio e instituciones de previsión. Nótese que hay un apartado específico de riqueza imponible, tema importante para el Negociado de Ingresos. Así pues este anuario cubre las dos funciones básicas por las que se crea el Negociado de Estadística: proveer información al Negociado de Ingresos y tener unas estadísticas presentables en el exterior.

Cabe destacar que la importancia relativa de los temas varía con el tiempo. Así, inicialmente las finanzas municipales ocupan casi el 20% de la información, mientras que diez años después su peso ha pasado a la mitad; en cambio los temas de educación y cultura juntamente con el de sanidad[5] adquieren importancia con el paso de los años (Urteaga y Nadal, 2002). Se puede hacer notar que el interés por datos de carácter social se incrementa a lo largo del periodo y en 1909 se crea el Museo Social de Barcelona.

La elaboración de estos anuarios que van firmados por el propio Manuel Escudé Bartolí son un logro muy importante. En su etapa en el Ayuntamiento de Barcelona, Escudé también se encargó de la aparición del *Boletín Municipal*, revista mensual que salió entre 1904 y 1906.

En el marco de su trabajo realizó estudios que cada vez se orientaban más claramente, quizá siguiendo la tendencia de la época, hacia temas de carácter social. En esta línea Urteaga y Nadal (2002) comentan que el “Censo obrero de 1905” publicado en el anuario sin nombre, lleva su firma. En 1921 aparece “Monografía estadística de las clases trabajadoras de Barcelona” de más de 100 páginas que ya había aparecido en el “Anuario Estadístico de la Ciudad de Barcelona” de 1917. Por otro lado, siguió conservando su perfil de geógrafo puesto que tras el “Atlas geográfico ibero-americano” hizo el volumen “Los Municipios de España” de 263 páginas (1901), que se reeditó en 1925 con los datos actualizados del censo de 1920.

En 1921 se crea el Instituto de Estadística i Política Social cuyo Presidente fue Manuel Escudé Bartolí. Sin embargo, tras su jubilación este desaparece puesto que el Ayuntamiento opta por la creación de un Instituto de Estadística Municipal que aglutina varias oficinas pero cuya dirección quedó desierta durante un periodo largo.

[5] Los datos de sanidad los facilitaba el Instituto de Higiene Urbana

Repercusiones de la obra de Manuel Escudé Bartolí

La obra de Manuel Escudé ha tenido una cierta repercusión tanto durante su vida, como años después. En vida, caben destacar dos aspectos: 1) su implicación en los temas relevantes del momento; así por ejemplo en el libro sobre Las Carolinas (1885), escrito en relación al interés del Imperio Alemán por tener colonias en el Pacífico, defiende los intereses colonialistas de España y el primer capítulo está dirigido al Gobierno (en concreto al Excmo. Sr. Presidente del Consejo de Ministros) y 2) su vertiente de geógrafo divulgador, ya se ha comentado que el “Atlas geogràfico de España” fue declarado de utilidad pública para la enseñanza.

Años después de su muerte, el gran legado de Manuel Escudé Bartolí son los “Anuarios Estadísticos de la Ciudad de Barcelona”. Massana (1971) señala que actualmente estos anuarios son una referencia imprescindible para conocer el detalle numérico de la ciudad. En este sentido pueden citarse numerosos trabajos que utilizan ya sean los datos o los análisis contenidos en los anuario, así por ejemplo en un estudio de genero Borderías (s. a.), hablando de las fuentes Tatjer (2001) o Porras (1995) desde una óptica epidemiológica.

Relación de trabajos por orden cronológico

1. (1878) - "Estudios sobre la historia física de la tierra" Barcelona (N. Ramírez) 1878, 25 pp.

2. (1880) - "Las maravillas del mar", 1880 [referenciado en EUI y en Olesti (1991)].

3. (1880) - "Diccionario estadístico español de los resultados generales del censo de la población, según el empadronamiento hecho en 31 de diciembre de 1877 y declarados oficiales por R.O. de 18 de abril de 1879" Barcelona (est. tip. Monrás y Bertran), 1880. 295 pp.

4. (1885) - "Las Carolinas. Descripción geográfica y estadística del archipiélago carolino. 1a parte; con datos recopilados y ampliados por Manuel Escudé Bartolí". Barcelona (impr. Luis Tasso Serra) 1885. 111 pp.

5. (1889) - "Memoria del estado actual de la agricultura, industria y comercio de la provincia de Barcelona", 1889 [referenciado en Alier (1974)].

6. (1891) - "España social y económica" (artículos publicados en el Diario Mercantil de Barcelona). Barcelona (impr. Diario Mercantil) 1891. 32 pp.

7. (1891) - "España y sus colonias: noticia de su población, agricultura, industria y comercio, según los más recientes datos estadçísticos" por Manuel Escudé Bartolí y Luis de Ramón, bajo la dirección de Pablo Riera. Barcelona (ed. de Riera) 1891. 277 pp.

8. (1895) - "La Producción española en el siglo XIX: estadística razonada y comparada". Barcelona (libr. A. Bastinos) 1895. 280 pp.

9. (1901)- "La natalidad de Barcelona", 1901 [referenciado en Alier (1974)].

10. (1901, 1904) - "Atlas geográfico de España", Barcelona (ed. A. Martín), 1901. 2 vols. 535 pp. + 52 hojas de láminas. Gran formato. Se publicó en fascículos, a partir del número 63 modifica el título "Atlas geográfico ibero-americano", los 65 primeros fascículos corresponden al Estado Español. A partir del 66 empieza Portugal. Se reedita en 1904 como "Atlas geogràfico ibero-americano"

11. (1901, 1925) - "Los municipios de España" Barcelona (ed. A. Martín) 1901. 263 pp. Se reedita en 1925 con los datos actualizados (datos del censo de 1920 y Estatuto Municipal de 1924), 267 pp.

13. (1921) - "Monografía estadística de la clase obrera" Barcelona (impr. de Henrich y Cia) 1921, pp. 128. Había sido publicado en el "Anuario Estadístico de la ciudad" de 1917.

Bibliografía

ALIER, R. (1974): "Escudé i Bartolí, Manuel", a "Gran Enciclopedia Catalana", vol. 6, p. 788. Gran Enciclopedia Catalana S.A.

ARQÜESO, A.; CELESTINO, F. (1999): "La estadística y la geografia vuelven a encontrarse". Revista Fuentes Estadísticas, vol. 38 [http://www.ine.es/revistas/fuentes/Numero38/pag4.htm (consultada 5/4/07)].

ARRÓ TRIAY, F. (1892) "Estadística médica de la Compañía de los Ferrocarriles de Tarragona a Barcelona y Francia correspondiente al septenio de 1879 a 1885 respecto de las líneas de Gerona y al año 1886 respecto de éstas y de las de Tarragona y la Frontera". Barcelona (impr. de Henrich y Cia en comandita Sucesores de N. Ramírez y Cia). Reimpreso en 1985 por el Ayuntamiento de Barcelona, Sociedad Catalana de Seguridad i Medicina del Trabajo y Universidad de Barcelona.

BORDERÍAS, C. (s.a.): "La evolución de la actividad femenina en la formación del mercado de trabajo barcelonés 1856-1930". Mimeo (Universidad de Barcelona).

CERDÀ, I. (1867): "Teoria general de la urbanización, y aplicación de sus principios y doctrinas a la reforma y ensanche de Barcelona", vol. II "La urbanización considerada como un hecho concreto. Estadística urbana de Barcelona". Madrid (impr. Española). Reimpreso en 1967 por el Instituto de Estudios Fiscales.

ENCICLOPEDIA UNIVERSAL ILUSTRADA EUROPEO AMERICANA, ESPASA-CALPE S.A. (s. a.), vol. 20, p. 1008.

FIGUEROLA, L. (1849): "Estadística de Barcelona en 1849". Barcelona (impr. y libr, Politécnica de T. Gorchs). Reimpreso en 1968 por el Instituto de Estudios Fiscales.

GRAS ELIAS, F. (1899): "Hijos ilustres de Reus". Barcelona (Libr. de Francisco Puig y Alfonso).

MASSANA, C. (1971): "Los anuarios estadísticos de la ciudad de Barcelona 1902-1923 y el socialismo municipal". Cuadernos de arquitectura y urbanismo, vol. 80, p. 47-54.

OLESTI TRILLES, J. (1991): "Diccionari biogràfic de reusencs", vol. 1, p. 228. Ajuntament de Reus.

PORRAS GALLO, Mª I. (1995): "La prensa madrileña de información general ante la epidemia de gripe de 1918-1919". Revista de estudios históricos de las ciencias médica, vol. 57.

RAFOLS, J. F. (dir.) (1980): "Diccionario de artistas de Cataluña, Valencia y Baleares". Ediciones Catalanas, S. A. y Gran Enciclopedia Vasca, vol 2, p. 374.

TATJER, M. (2001): “Estadísticas de Barcelona 1841-1960”. Revista Fuentes Estadísticas, vol. 50. [http://www.ine.es/revistas/fuentes/Numero50/paginas/22.htm (consultada 5/4/07)].

URTEAGA, L.; NADAL, F. (2002): “L’organització del servei d’estadísitca de l’Ajuntament de Barcelona (1902-1923)”, a VENTURA, A. (dir.): “100 Anys d’estadísitca municipal”. Ajuntament de Barcelona. p. 21-52.

VENTURA, A. (dir.) (2002): “100 Anys d’estadísitca municipal”. Ajuntament de Barcelona.

Capítulo 28

Orígenes de la bioestadística en España: Estadísticas demográficas y sanitarias

ANTONIO FRANCO RODRIGUEZ-LAZARO
MERCEDES CASAS GUILLEN
Universidad CEU San Pablo de Madrid

Introducción

La bioestadística en España tiene sus orígenes en la elaboración de estadísticas demográficas sanitarias, a mediados del siglo XVIII. A lo largo de los siglos XIX – XX se constituyen organismos oficiales que registran la información estadística de salud demográfica y se editan normas generales de salud pública.

Paradigmas de la Bioestadística. Estadísticos europeos

Siguiendo a los autores José Almenara y José Carlos Silva[1] se pueden identificar dos paradigmas clave en la historia de la Bioestadística que han permitido su evolución hasta la ciencia actual. El primer paradigma es denominado por los autores como *aritmético, político y social*, y constituye la génesis de la bioestadística. Se presenta en el siglo XVII cuando se independiza la Estadística de la mera actividad censal para desarrollarse como corriente científica. Las figuras más destacables de este período son los ingleses John Graunt y William Petty, o franceses tales como Philippe Pinel y Pierre Charles Alexandre Louis. Se produce un cambio en la forma de abordar problemas tradicionales como la periodicidad de aparición de ciertas enfermedades, el conocimiento de las crisis de mortalidad, la observación de la regularidad de sexos en los nacimientos, etc. Esta etapa perdura hasta la época de la Revolución Francesa, cuando se utiliza la aritmética como una herramienta política y social.

El segundo paradigma de la bioestadística representa el origen de la estadística probabilística e inferencial moderna por lo que se denomina *inferencial y biomédico*. Está

[1] ALMENARA, J; SILVA, L.C. (2003) *Historia de la bioestadística: génesis, la normalidad y la crisis*. Quórum Editores, Cádiz. Pág. 29

basado en el concepto de normalidad y se consolida en la primera mitad del siglo XX con el desarrollo de la estadística inferencial. Los científicos más representativos de esta época son: Francis Galton, Karl Pearson, Major Greenwood, Raymond Pearl o Wade H. Frost.

Comentaremos brevemente lo más destacable de los científicos mencionados, y posteriormente analizaremos con detalle el desarrollo de la Bioestadística en España.

John Graunt (1620–1674) fue un comerciante de tejidos que se interesó por la información demográfica contenida en los Boletines de Mortalidad de la ciudad de Londres. Elaboró una Tabla recopilatoria para una población estacionaria, intentando estimar la tasa anual de mortalidad, identificando ciertas regularidades y patrones en algunas enfermedades de la población, y calculando la frecuencia de algunas variables, como por ejemplo que era mayor la proporción de niños nacidos vivos que la de niñas. Puede considerarse que *Natural and Political Observations mentioned in a following index, and made upon the bills of mortality. By John Graunt, citizen of London. UIT reference to the government, religión, trade, growth, ayre, diseases, and several changes in the said city (1662)* es el primer trabajo cuantitativo que interpreta las características de la conducta social. Es además uno de los trabajos fundamentales en la confección de tablas de esperanza de vida.

Un economista, músico, científico, poeta, estadístico y médico, amigo de Graunt, llamado Sir William Petty (1623–1687)[2], realizó importantes investigaciones en estadística demográfica publicando trabajos relacionados con los patrones de mortalidad, natalidad y enfermedad de la población inglesa. Petty propuso la creación de una administración estatal que analizase los movimientos de la población, así como sugirió la elaboración de tablas de mortalidad clasificadas por edad de ocurrencia, anticipándose al desarrollo de las actuales tablas usadas para comparar poblaciones diferentes. Se interesó en sus investigaciones por aquellos enfermos que se curaban de sus dolencias por azar, aunque hubiesen recibido atención médica para remediarlas La metodología de tratamiento de la información poblacional fue denominada por Petty "aritmética política": *Political Arithmetick (publicada en 1690 y Political Anatomy of Ireland (publicada en 1691).*

El médico francés Philippe Pinel (1745–1820) estableció la primera escuela de psiquiatría en su país, adoptando la observación empírica como metodología de investigación. Conoció los trabajos de los científicos ingleses cuando colaboraba con la "Royal Society de Londres", la primera asociación de estadísticos creada en 1834. En 1809 Pinel publicó la segunda edición de *Traité médico-philosophique sur l'aliénation mentale*, donde introduce el cálculo de probabilidades en la clasificación de los enfermos mentales atendiendo a las causas de las patologías[3].

Pierre Charles Alexandre Louis (1787 – 1872) fue uno de los precursores del llamado "método numérico" en Medicina. Esta técnica empleaba gran cantidad de información cuantitativa y cualitativa de los pacientes para establecer conclusiones acerca de las enfermedades poblacionales. Durante los siete años que trabajó en el Hospital de la Charité recopiló mucha información de sus pacientes sobre antecedentes familiares, evolución de sus dolencias, diferentes tratamientos aplicados, resultados positivos y negativos obtenidos, etc. Louis planteó la posibilidad de que los sucesos epidemiológicos se comporten siguiendo leyes matemáticas similares a las que rigen los fenómenos naturales. En uno de sus trabajos más

[2] http://escuela.med.puc.cl/Recursos/recepidem/introductorios3.htm

[3] PINEL, P. *Traité médico-philosophique sur l'aliénation mentale* (1995), pág. 436. Publicación electrónica: http://gallica.bnf.fr/ark:/12148/bpt6k76576g

conocidos demuestra que la tuberculosis no se transmite hereditariamente y pone en tela de juicio la utilización de la sangría en enfermedades pulmonares[4].

El método numérico de Louis influyó en los epidemiólogos de los siglos XIX-XX de tal modo que se interesaron en encontrar regularidades matemáticas en las enfermedades de las poblaciones. No obstante, la ciencia epidemiológica tiene como premisa que las enfermedades no ocurren al azar, por lo que sus investigaciones se encaminan a identificar, con la mayor precisión posible, las realidades que pueden ser calificadas como "causas" de las enfermedades diferenciándolas de aquellas que aparecen de forma aleatoria. Los resultados de las investigaciones llevaron a la aplicación de las leyes de los grandes números al análisis de las estadísticas sanitarias y a la creación de conceptos fundamentales en la teoría epidemiológica como "tasa estandarizada", "medición año–persona" y "exposición poblacional"[5].

El estudio de los registros poblacionales efectuado por las autoridades de los diferentes países europeos originó y fundamentó la modelización estadística de la Epidemiología, sobre todo a partir de los trabajos de William Farr[6] (1807–1883), Marc d´Espine (1806–1860) y Jacques Bertillon (1851–1922). El epidemiólogo y estadístico inglés William Farr estableció un sistema de recogida rutinaria de datos y la práctica de utilizar estos datos para evaluar problemas sanitarios, a partir de la información obtenida de los censos realizados en Inglaterra entre 1801 y 1831.

Algunos años después, la Oficina del Registro General de Inglaterra y Gales encargó a William Farr un estudio encaminado a conseguir una terminología internacional de las causas de morbilidad y mortalidad con la que realizar estudios estadísticos. El doctor Farr especificó una primera nomenclatura estadística sobre las causas de defunción que tuvo una gran aceptación en el primer Congreso Internacional de Estadística (Bruselas, 1853). Una segunda enumeración fue elaborada por el doctor Farr y el médico suizo Marc D'Espine, que consistía en un listado de las causas de morbilidad que aparecían con mayor asiduidad. Finalmente, la nomenclatura de causas de defunción definitiva fue realizada por el estadístico francés Jacques Bertillon[7 y 8] (1851–1922), que fue ratificada por el Instituto Internacional de Estadística en la reunión celebrada en 1893 en Chicago. En 1900 se celebró en París la Primera Conferencia Internacional para la Revisión de la clasificación internacional de enfermedades de Bertillon.

---

[4] LOUIS, PIERRE C.A. (1835) *Recherches sur les effets de la saignée dans quelques maladies inflammatoires et sur l'action de l'émétique et des vésicatoires dans la pneumonie*. Pág. 1. Publicación electrónica: http://gallica.bnf.fr/ark:/12148/bpt6k103986t.chemindefer

[5] LÓPEZ MORENO, S. (1998) "Acerca del Estatuto Científico de la Epidemiología". *Revista Salud Pública de México, septiembre-octubre, vol.40, núm. 5*. Instituto Nacional de Salud Pública, Cuernavaca (México), págs. 389 - 391. Pub. electrónica: http://redalyc.uaemex.mx/redalyc/pdf/106/10640501.pdf

[6] EYLER, J. M. (1979) *Victorian social medicine. The ideas and methods of William Farr*. Baltimore. The Johns Hopkins University Press.

[7] BERTILLON, J. (1899) *Nomenclatura de las enfermedades. Causas de defunción – Causas de incapacidad para el trabajo*. Adoptada por el Servicio de Estadística de la Ciudad de París. Traducción española de la Secretaría del Consejo Superior de Salubridad de México. Págs. 1 – 7.

[8] En la obra de Jacques Bertillon ya citada, la clasificación de las enfermedades es la siguiente:
I.- Enfermedades Generales; II.- Enfermedades del sistema nervioso y de los órganos de los sentidos.; III.- Enfermedades del aparato circulatorio; IV.- Enfermedades del aparato respiratorio; V.- Enfermedades del aparato digestivo; VI.- Enfermedades del aparato génito – urinario y de sus anexos; VII.- Enfermedades puerperales; VIII.- Enfermedades de la piel y de sus anexos; IX.- Enfermedades de los órganos de la locomoción; X.- Vicios de conformación; XI.- Enfermedades de la primera infancia; XII.- Enfermedades de la vejez; XIII.- Afecciones producidas por causas exteriores; XIV.- Enfermedades mal definidas.

Los impulsores de la epidemiología estadística del siglo XX en Inglaterra y Estados Unidos, respectivamente, fueron Major Greenwood (1880–1949) y Raymond Pearl (1879–1940). Ambos fueron discípulos y seguidores inicialmente de la escuela "eugenesia[9]" que inició Francis Galton[10] (1822–1911) y terminaron siendo críticos con dicha disciplina. Tanto Greenwood como Pearl se interesaron por la Biometría ampliando su formación matemática con Karl Pearson (1857–1936). El laboratorio Biométrico dirigido por Karl Pearson[11] (1857–1936) utilizó la Estadística en la búsqueda de respuestas a los hallazgos obtenidos en la investigación de diferentes problemas biológicos, en especial los derivados de la eugenesia.

Greenwood (1880–1949) fue catedrático en epidemiología y estadística demográfica en la London School of Higiene and Tropical Medicine desde 1927 hasta 1938, difundiendo sus enseñanzas en un tratado clásico *Epidemics and Crowd Diseases. An Introduction to the Study of Epidemiology*[12]. Fue un gran defensor de la metodología experimental, siendo su trabajo más conocido *Experimental Epidemiology*[13] sobre la búsqueda de un modelo estadístico que reflejase el comportamiento durante quince años de una enfermedad infecciosa que afectaba a una comunidad de ratones.

El norteamericano Wade Hampton Frost (1880-1944), epidemiólogo contemporáneo de Greenwood, ejerció también como profesor en la Escuela de Salud Pública de la Universidad Johns Hopkins. Greenwood y Frost consideraban que la epidemiología era el instrumento que promovía dos líneas diferentes de investigación; la primera era el análisis durante periodos de tiempo amplios de la distribución de las enfermedades y de los determinantes que afecten al colectivo concreto que estamos estudiando, por lo tanto, la epidemiología queda desvinculada de la práctica clínica. La segunda línea investigadora consistía en el descubrimiento de la proporción de afectados por una epidemia o endemia en una localidad determinada.

El biólogo norteamericano Raymond Pearl (1879–1940) estudió medicina en Alemania, pero continuó su formación estadística en Inglaterra y Estados Unidos, donde fue profesor de biometría y estadísticas demográficas en la Escuela de Salud Pública de la Universidad Johns Hopkins (1918-1925), pasando posteriormente a desempeñar el puesto de estadístico de su hospital hasta 1930. Sus investigaciones sobre la dinámica demográfica le permitieron elaborar la curva de crecimiento exponencial de la población, así como estimar y predecir cambios en la población mundial.

Los avances en la elaboración de diseños experimentales, los nuevos descubrimientos en Física y en Química, así como el hallazgo de las causas de muchas enfermedades debido a la existencia de pruebas bacteriológicas, hizo que existiera una mayor colaboración entre la Estadística y la Biología al analizar la dinámica y genética de las poblaciones y en la aplicación de la Estadística en el origen y evolución de las epidemias[14]. En la segunda mitad del siglo XX la Epidemiología experimenta un gran desarrollo al sustituir el interés por desarrollar modelos que reflejaban la evolución de las epidemias, línea de investigación fomentada por los trabajos de William Farr; por estudios que analizaban la información disponible de las enfermedades crónicas. En ese contexto, lo que pretende la epidemiología es

---

[9] Eugenesia: aplicación de las leyes biológicas de la herencia al perfeccionamiento de la especie humana.

[10] ÁLVAREZ PELÁEZ, R. (1990) *Sir Francis Galton, padre de la Eugenesia*. Madrid. CSIC.

[11] PEARSON, E.S. (1948) *Pearson, creador de la estadística aplicada*. Buenos Aires-México. Espasa Calpe Argentina S.A.

[12] GREENWOOD, M. (1935) *Epidemics and Crowd Diseases. An Introduction to the Study of Epidemiology*. London. Williams & Nogate.

[13] GREENWOOD, M. BRADFORD HILL, A. TOPLEY, WWC. Y WILSON, J. (1936) *Experimental Epidemiology*. Medical Research Council Special Report, series nº209. London HMSO 1936.

[14] ISRAEL, G. (1993) *The emergente of biomathematics and the case of population dynamics. A revival of the mechanical reductionism and Darwinism*. Science in Context. 6 (2) 469-509.

descubrir los factores de riesgo o relaciones causales y, en su caso, el agente causal único de cada enfermedad, encontrándose con la dificultad adicional de que no es fácil ver la relación causa-efecto en las enfermedades crónicas.

Estadísticas demográficas y sanitarias elaboradas con la información censal. (S. XVI – XVIII)

Desde el siglo XVI encontramos en España una actividad censal muy productiva, aunque no exhaustiva en todos los Reinos de España, como expone Manuel Martín[15]:

> *En Castilla se practicó una amplia y minuciosa encuesta para el reparto de «servicios» entre 1528 y 1536, y se elaboraron unos padrones para el reparto de «millones» en 1591 […]. En Vascongadas se dispone de varios recuentos en cada una de las provincias, aunque con datos incompletos. Tomás González[16] publicó dos censos referentes a Navarra, de 1553 y 1587, y en el Archivo del Reino de Navarra existen apeos de 1637, 1646 y 1677 para el reparto de «cuarteles y alcabalas». Para los territorios de la Corona de Aragón las fuentes son más escasas, pero existen recuentos parciales para Aragón en 1603 y 1650, para Cataluña en 1515 y 1553, y para Valencia en 1562 y 1609 […].*

En el siglo XVII se produce un período de carencia de datos estadísticos que Manuel Martín Rodríguez imputa a la Administración española: *la debilidad de la Administración de los últimos Austrias queda bien patente en el vacío informativo demográfico del siglo XVII, en el que no se puede contar ni con un solo recuento completo […]. Sólo quedan, por tanto, para este siglo, los registros municipales parroquiales.*

Es en el siglo XVIII cuando se elaboran los censos considerados como las primeras fuentes estadísticas españolas de trascendencia:

- El Catastro de Jerónimo de Uztariz corrigió el vecindario de Campoflorido realizado entre 1712 y 1717.
- El Catastro del Reino de Castilla efectuado por el Marqués de la Ensenada entre 1749 y 1753 supone la culminación de una intensa actividad estadística iniciada en la década de los cuarenta
- El Censo de Aranda (1768 – 1769) constituye un intento riguroso de aprovechar con fines civiles la extensa información demográfica contenida en los archivos parroquiales
- Los censos de Floridablanca (1787) y de Godoy (1797) marcan una pauta diferente en el análisis de los datos, puesto que presentan clasificaciones cruzadas de la población según sexo, intervalos de edad, estado civil u ocupación, por zonas geográficas españolas.

Paralelamente a la actividad censal se desarrolla una importante escuela de probabilistas en España, de la que es precursor Juan Caramuel, que se extenderá al resto de Europa en el ámbito de la teología moral[17].

---

[15] MARTÍN RODRÍGUEZ, M. (1984) *Pensamiento económico español sobre la población*. Pirámide, Madrid. Págs. 29-31.
[16] MANUEL MARTÍN RODRÍGUEZ CITA A: GONZÁLEZ, T. (1829) *Censo de población de las provincias y partidos de la corona de Castilla en el siglo XVI*, Madrid.
[17] MARTÍN PLIEGO, F.J (2002) "Los probabilistas españoles de los S. XVII a XIX". *I Jornadas de Historia de la Probabilidad y de la Estadística*, AHEPE. AC, Madrid, Pág.

En los últimos años del siglo XVIII, bajo el reinado de Carlos III, la Corona comienza a interesarse por preservar la salud pública en España, redactándose normas sobre higiene, como por ejemplo la obligación de trasladar los cementerios fuera de los núcleos urbanos[18] y el establecimiento de un sistema de saneamiento con colectores en las ciudades que conducían las aguas fecales hacia depósitos situados en los extrarradios, incrementando las exigencias sanitarias en las viviendas de los habitantes. En esa época se trasladó la competencia que tenía la Iglesia sobre el control de la población a organismos de naturaleza civil, en lo que respecta a nacimientos, defunciones, matrimonios, etc, creándose a tal efecto los registros civiles. No obstante, a finales del siglo XIX todavía se encontraban nacimientos y matrimonios registrados en parroquias que no se hallaban reflejados en los registros civiles de las grandes ciudades[19].

La organización de las Estadísticas modernas. Primeros Anuarios Estadísticos. (S.XIX)

Gracias al desarrollo de la estadística, la historia de la demografía y la epidemiología disponía de nuevos recursos de investigación que dieron lugar a la elaboración de Tablas de mortalidad, natalidad, así como al estudio del movimiento natural de la población a partir de los datos obtenidos de los Registros Civiles.

El marino, poeta satírico y matemático ilustrado José Vargas Ponce (1760 – 1821)[20] realizó la estadística de Guipúzcoa y unas tablas de mortalidad, similares a las de John Graunt[21], de todos sus pueblos desde 1701 a 1800 como fruto de las observaciones realizadas en sus viajes a esta zona. El trabajo fue publicado en 1805 con el nombre de *Estados de Vitalidad y Mortalidad de Guipúzcoa en el siglo XVIII*. Era un gran conocedor de la aritmética de su tiempo, lo que le permitió elaborar un largo tratado explicando todos los cálculos de la aritmética mercantil y "quintas reglas" que se utilizaban en la práctica mercantil: "cambios", "días fijos de los pagos", etc.

La organización de las estadísticas modernas en España comienza en 1856 con la creación de la *Comisión Estadística General del Reino*, que en 1861 pasa a denominarse *Junta General de Estadística del Reino* y en 1873 fue reemplazada por el *Instituto Geográfico y Estadístico*. Estas instituciones confeccionaron los primeros Anuarios Estadísticos y estudiaron la población de España. En 1858 aparecen las primeras estadísticas oficiales del movimiento natural de la población española, que se elaboraron con los datos recogidos en los archivos parroquiales. Tal y como se expone en la *Guía de fuentes cuantitativas para la historia económica de España (I)*, no se puede precisar cuándo comenzaron a utilizarse estos registros parroquiales, pero lo que no cabe duda es que sin ellos no hubiera sido posible elaborar ningún estudio demográfico. Sin embargo, estos libros no estaban exentos de errores, existía una tendencia a la infravaloración de la natalidad en los registros de bautismos porque no se solía anotar los nacidos muertos[22].

---

[18] Real Cédula de 13 de abril de 1787, "Novísima Recopilación". Véase: M. MARTÍN RODRÍGUEZ. *Pensamiento económico español sobre la población.*, pág. 268.

[19] "Movimiento de la Población de España" (Madrid 1895), 11 y 25. Véase: RODRÍGUEZ OCAÑA, E. Y BERNABEU MESTRE, J. (1997*). Physicians and statisticians: two ways of creating demographic health statistics in Spain, 1841 – 1936*. *Continuity and Change 12 (2*), pág. 249.

[20] http://www.cervantesvirtual.com/servlet/SirveObras/01472842099194951154480/p0000001.htm

[21] ALMENARA, J; SILVA, L.C. (2003) *Historia de la bioestadística: génesis, la normalidad y la crisis*. Quórum Editores, Cádiz. Pág. 151.

[22] COLL, S.; FORTEA, J.I. (1995) *Guía de fuentes cuantitativas para la historia económica de España*, Vol. I Recursos y sectores productivos. Banco de España, Servicio de Estudios, 1995. Pág. 47

Los Anuarios Estadísticos contenían información de los centros de asistencia sanitaria existentes: hospitales provinciales y municipales, hospicios, asilos, casas de maternidad y centros de atención a la infancia. Recogían cifras de enfermos atendidos, número de partos y mujeres y/o niños fallecidos en dichos partos, defunciones y algunas causas de defunciones. Se mostraba la información agrupándola según las provincias o capitales de provincia. Asimismo, existía un anexo donde se comparaban las defunciones por edad de España y las de otros países, y una sección adicional con información de los datos climáticos de las regiones.

Desarrollo de las instituciones estadísticas de salud pública. Higienistas y epidemiólogos españoles (S.XIX- XX)

El siglo XIX marca el inicio del desarrollo de una extensa normativa reguladora de higiene y salud pública, en la que intervinieron, además del gobierno, médicos e intelectuales ilustrados.

En 1899 y en 1902 se producen dos hitos importantes en el tratamiento de la información estadística en España: adopción de la Normativa Internacional de Clasificación de Enfermedades y Causas de la Muerte (*Classification of Diseases and Causes of Death*) por el Instituto Geográfico y Estadístico dada por Bertillon, y la publicación de las series anuales de estadísticas demográficas por dicho organismo.

El 17 de marzo de 1847 se crea el *Real Consejo de Sanidad* en virtud del Real Decreto Orgánico de Sanidad, desarrollado por Real Orden del 26 de marzo sobre “el reglamento organizativo y atribuciones del Consejo y las Juntas de Sanidad”. Estas normas configuran el antecedente de la Ley Orgánica de Sanidad del 28 de noviembre de 1855, que incluye la creación de la *Dirección General de Sanidad* como órgano ejecutivo. En el organigrama se preveía la necesaria inclusión de médicos “higienistas” de reconocido prestigio y dispuestos a prestar su colaboración a la Administración de forma gratuita.

El gran desarrollo de la doctrina higienista en España se debe, en gran medida, a dos importantes médicos, el internista Ignacio María Ruiz de Luzuriaga, que ejerció durante varios años en escuelas británicas, y posteriormente regresó a España para dirigir acciones de higiene pública; y el científico Mateo Seoane Seobral, que también tuvo contacto con otros científicos británicos y como docente ejerció una influencia decisiva en sus discípulos, entre los que destacan especialmente Pedro Felipe Monlau y Francisco Méndez Álvaro.

Mateo Seoane Seobral (1791–1870) estudió Medicina inicialmente en Valladolid, su ciudad natal, y posteriormente en Salamanca, doctorándose en 1813. Su brillante carrera como médico rivalizaba con su actividad política de ideas liberales. Colaboró en la elaboración de la Primera Ley de Beneficencia, y siendo diputado presentó el Proyecto de Reglamento General de Sanidad de 1822. Este proyecto no llegó a ponerse en práctica y hay que esperar a mediados del siglo XIX para que se adoptaran medidas preventivas de higiene y salud pública, encaminadas fundamentalmente a evitar el contagio de enfermedades de transmisión sexual como la sífilis, estableciendo controles médicos periódicos de las prostitutas. Mateo Seoane fue perseguido en España por sus ideas políticas teniendo que exiliarse a Londres, en donde colaboró con importantes médicos ingleses pertenecientes al Colegio de Médicos de la Sociedad Médica de Londres y al Instituto Real de la Gran Bretaña.

La llegada de una epidemia de cólera a España hizo que las autoridades sanitarias requirieran médicos expertos en esta enfermedad al gobierno inglés para formar parte de la Junta Superior Gubernativa de Medicina y Cirugía de España. Paradójicamente el gobierno inglés recomendó al Doctor Seoane, por el prestigio y reconocimiento obtenido en su país.

Las condiciones impuestas por Mateo Seoane a su regreso fueron que la Junta no censurara sus escritos a cambio de no publicarlos[23]. Obras destacables de Mateo Seoane en el ámbito sanitario son el *Informe acerca de los principales fenómenos observados en la propagación del cólera indiano por Inglaterra y Escocia en 1832* y las *Instrucciones generales sobre el modo de preservarse del cólera morbo epidémico en 1834.*

Su conocimiento de la Estadística queda reflejado en el ensayo *Consideraciones generales sobre la estadística médica (1838)*[24]. En esta obra explica su utilización en las últimas décadas del siglo XIX tanto en España como en otros países europeos, especificando las causas de la incorrecta aplicación de su metodología en la Medicina. Algunos médicos se dedicaban a recopilar datos sin reflejar los antecedentes familiares y las circunstancias socioeconómicas de los pacientes, omisión que podría ejercer influencia sobre sus diagnósticos. Otros médicos actuaban con criterios opuestos, teniendo en cuenta únicamente la doctrina tradicional en lugar de recoger información empírica con la que validar o anular sus conjeturas.

El doctor Seoane considera que es inadecuada la costumbre que tienen los médicos de generalizar los resultados de sus investigaciones sin haber reunido un gran número de casos de forma correcta y ordenada, o por recoger la información de forma deficiente, lo que puede provocar errores en la estimación del alcance de las enfermedades y/o epidemias.

En el año 1843 se crearon en España las dos primeras cátedras de Estadística, aunque no tuvieron carácter universitario[25], y de forma simultánea aparecía la asignatura higiene pública en los planes de estudio de la carrera de Medicina. El manual utilizado para impartir docencia es *Elementos de Higiene Pública* de Pedro Felipe Monlau (1808–1871) publicado en 1847 y reeditado posteriormente en 1862 y1871. Monlau plantea unas reflexiones generales sobre la estadística, utilizando argumentos similares a los que Seoane había expuesto treinta años antes que fueron adoptados por la mayoría de los profesores españoles de higiene pública antes de la Segunda República (1931). Asimismo, el doctor Monlau alude al fundamento y objeto de las dos estadísticas que con una perspectiva global se efectuaban en ese momento, la estadística administrativa de cada país, incluyendo territorio y población, y la estadística médica, también denominada higiénica o sanitaria. La estadística médica agregaba datos de edad, sexo, profesión y hábitos, clase social, características de la habitación y del régimen de vivienda, temperamento y constitución general, constitución médica, constitución epidémica, tipologías hereditarias que producen enfermedades, diátesis (predisposición orgánica a contraer una enfermedad determinada), enfermedades padecidas, causas que influían de forma determinante en las defunciones, número de autopsias realizadas, clima y topografía[26]

El conflicto entre los fundamentos de la experimentación en laboratorio y los principios socio-morales basados en la cuantificación estadística queda reflejado en la discusión que mantuvieron el inspector provincial de Madrid José Call y Morros (1858-c.1923) y José Nin y

---

[23] VIÑES RUEDA, J.J. (2006). *La Sanidad española en el siglo XIX a través de la junta provincial de sanidad de Navarra 1870 – 1902*. Colección: Temas de historia de la medicina. Gobierno de Navarra, Departamento de Salud, 2006. Pág. 740

[24] SEOANE SOBRAL, M. (1838) *Consideraciones generales sobre la Estadística Médica*. Memoria leída en la sección de Ciencias Antropológicas de la Real Academia de Ciencias Naturales. Madrid, Imprenta de la compañía tipográfica, 1838. Págs. 1 – 34.

[25] SÁNCHEZ-LAFUENTE FERNÁNDEZ, J. (1975) *Historia de la estadística como ciencia en España 1500-1900*. Madrid. INE.

[26] SANTERO, F.J. (1885) *Elementos de Higiene privada y pública*. Madrid. El Cosmos Ed. P.168.

Pullés (+1892)27. El higienista madrileño José Call se oponía a la iniciativa de realizar estadísticas que tuvieran como finalidad demostrar que la distribución de la mortalidad era desigual por razones socioambientales, realidad que Nin denominaba "malaria urbana", ya que Call consideraba que esas mediciones tenían un carácter limitado por lo que era imposible que pudieran proporcionar información suficiente para afirmar con rotundidad la existencia de este tipo de relación. El barcelonés Nin, por el contrario, creía que el factor socioambiental era determinante a la hora de estudiar las posibles causas de los fallecimientos, teoría que expuso en la ponencia *Influencia que el modo de ser de las grandes urbes ejerce en la salud y en la longevidad de sus habitantes; aplicación de este estudio a nuestra ciudad,* presentada en el congreso médico celebrado en Barcelona en 1888[28].

El sector de los seguros comenzó a desarrollarse en España en el siglo XIX con una tabla de vida semejante a las calculadas por Quetelet realizada en 1866 por el matemático Miguel Merino (1831–1904) que incluía datos de enfermedades, tanto las que eran mortales como las causantes de invalidez. El estudio se realizó en Cataluña durante cinco años recogiendo información a través de sociedades mutualistas relativa 16.000 personas. Consideraba la estabilidad de la mortalidad en toda la población y un reparto homogéneo de la misma por grupos de edad. En 1868 ingresa como numerario en la Real Academia de Ciencias Exactas, Físicas y Naturales con el discurso *Del origen, importancia y aplicaciones del cálculo de las probabilidades*.

Francisco Méndez Álvaro (1806–1883) nació en Pajares de Adaja, un pueblo de Ávila. Desde muy joven recibió la influencia de sus tíos, Pedro y Aniceto de Álvaro, que eran de ideología liberal. En 1823 se matriculó en el Colegio de cirugía San Carlos de Madrid, ingresando en 1836 en la Plana Mayor del Cuerpo de Sanidad Militar, dirigido por el doctor Mateo Seoane. Decidió abandonar el ejercicio de la medicina para dedicarse a la actividad sociosanitaria, a la política y sobre todo al periodismo médico-político. En 1854, la revista por él dirigida "Boletín de Medicina y Cirugía" y la "Gaceta Médica" fundada por el doctor Matías Nieto y Serrano desaparecen con la finalidad de crear una nueva publicación dirigida por Méndez Álvaro, denominada "El Siglo Médico", que dejó de publicarse en 1936. Al igual que el doctor Seoane, Méndez Álvaro colaboró en los informes previos a la promulgación de la Ley de Sanidad de 1855, así como en el dictamen de medidas preventivas de salud e higiene contra la lepra en el año 1878.

Desde su posición como editor del periódico El Siglo Médico, impulsó la creación de Instituciones estadísticas centrales y provinciales en las que debían participar profesionales médicos. Era partidario de establecer una red comunicativa entre la Real Academia de Medicina de Madrid, el Consejo Real de la Salud y los organismos estadísticos provinciales, así como de unificar las administraciones sanitarias y estadísticas[29].

A pesar de las regulaciones establecidas en este período, seguían existiendo deficiencias en la recogida de estadísticas sanitarias, así como en su difusión a través de algunos medios informativos (*Gaceta de Madrid*[30]), por lo que el objetivo de establecer una estructura

---

[27] NIN Y PULLÉS, J. (1888) *Influencia que el modo de ser de las grandes urbes ejerce en la salud y en la longevidad de sus habitantes; aplicación de este estudio a nuestra ciudad.* Gaceta Sanitaria de Barcelona. 1, 114-120.

[28] *Congreso de Ciencias Médicas de Barcelona…1888* (1889) Barcelona. Exposición Universal, p. 935-950.

[29] "Movimiento de la Población de España" (Madrid 1895), 11 y 25. Véase: RODRÍGUEZ OCAÑA, E. Y BERNABEU MESTRE, J. (1997*). Physicians and statisticians: two ways of creating demographic health statistics in Spain, 1841 – 1936. Continuity and Change 12 (2),* pág. 250.

[30] "La Gaceta de Madrid" es el documento público antecesor del actual Boletín Oficial del Estado (BOE). Véase: http://www.boe.es/g/es/bases_datos/gazeta_ayuda.php

organizativa homogénea y continua de Estadísticas de Salud en España aún a finales del siglo XIX estaba muy lejano. Tal y como explican Rodríguez Ocaña y Bernabeu Mestre: *desde nuestra actual perspectiva, es evidente que la falta de experiencia técnica, y la poca atención prestada a los estudios estadísticos dentro de la medicina y a los estudios de universidades en general, contribuyeron al retraso en la adquisición y explotación de registros fiables*[31].

En 1888, aparece el nuevo *Boletín de Sanidad* de periodicidad anual, que incorporaba información específica de Madrid aunque carecía de datos estadísticos de otras ciudades españolas. Poco después se vuelve a producir un período de ausencia de publicaciones estadísticas en España, y existen varias disposiciones reguladoras en las que se delega la responsabilidad de obtener información a los gobiernos municipales.

La reorganización sanitaria en España. Primer tercio del siglo XX

A principios del siglo XX se producen ciertos acontecimientos que llevan a un proceso de modernización[32] de la sociedad española. La independencia de las colonias americanas supone una disminución en los recursos materiales y de mano de obra, así como una reducción en las arcas españolas. La economía española tiene que reestructurarse para continuar creciendo, y la Administración debía reorganizarse en todos los sectores, incluido el de la Sanidad Pública.

El comienzo de la reforma de los servicios sanitarios llega con el *Real Decreto General de Sanidad de 12 de enero de 1904*, que reforma la Ley de Sanidad vigente desde 1855. Esta normativa extiende la organización sanitaria hasta el nivel local delegando competencias sanitarias en los municipios españoles, así como obligando a los ayuntamientos a dictar normas de higiene pública. Por este motivo la Comisión de Estadísticas del Consejo Real de Sanidad incluye en el *Boletín Estadístico Sanitario* la publicación semestral de datos recibidos de los funcionarios provinciales de sanidad, a partir de julio de 1905.

Mediante la *Real Orden del 2 de julio de 1909* se produce una reestructuración del servicio de estadísticas sanitarias por la que se ordena la publicación mensual de datos de las capitales de provincia y las ciudades de más de 10.000 habitantes. Dicha información debía contener datos generales de mortalidad, muertes por enfermedades específicas y contagiosas, morbilidad en hospitales e instituciones de caridad famosas, nacimientos, variables meteorológicas y estadísticas especiales.

En 1910 aparece por primera vez la Epidemiología como disciplina profesional tras la reforma realizada en el Instituto Alfonso XIII en la que se crea una sección dedicada a esta ciencia. En sus comienzos, esta sección se dedicaba sobre todo a luchar contra focos epidémicos de enfermedades contagiosas, antes que a innovar y desarrollar investigaciones sanitarias[33].

Nuevamente se edita el *Boletín de Estadística Demográfico-Sanitaria*[34], con periodicidad mensual, que se mantendrá hasta 1920 cuando la *Real Orden del 26 de Noviembre de 1920*

[31] "Movimiento de la Población de España" (Madrid 1895), 11 y 25. Véase: RODRÍGUEZ OCAÑA, E. Y BERNABEU MESTRE, J. (1997*). Physicians and statisticians: two ways of creating demographic health statistics in Spain, 1841 – 1936*. *Continuity and Change 12 (2*), pág. 251.

[32] MARTÍNEZ NAVARRO, J.F. (1994) "Salud pública y desarrollo de la epidemiología en la España del siglo XX". *Revista de Sanidad e Higiene Pública, vol. 68*, monográfico. Págs.29 – 43.

[33] MARTÍNEZ NAVARRO, F. (1992) "Salud Pública y desarrollo de la epidemiología en la España del siglo XX". *Revista de Sanidad e Higiene Pública*, vol. 68, Págs. 29-43.

[34] *Boletín mensual de Estadística Demográfica-Sanitaria del mes de Diciembre de 1911*. Págs. 611 – 666.

termina con la publicación de este boletín y permite la edición del *Anuario de la Dirección General de Sanidad*. Este anuario recogía temas de interés de la sanidad pública en dos bloques temáticos: uno de acciones técnicas, jurídicas y administrativas relativas a sanidad y otro bloque estadístico.

En la dictadura de Primo de Rivera, se constituyen importantes organismos sanitarios, como la Escuela de Sanidad Pública creada en 1924, que será un centro de formación bioestadística de los profesionales sanitarios españoles. Por otro lado, se concedía menos importancia a la difusión de estadísticas sanitarias, por lo que el Boletín Técnico de la Dirección General de Sanidad solamente empleaba los datos relativos a nacimientos y defunciones que recopilaba el Instituto Geográfico y Estadístico.

La Ley sanitaria local de 1925 transfiere más competencias sanitarias a los responsables provinciales y municipales con el respaldo político de los Gobernadores y Alcaldes. Esta concordancia entre las instituciones sanitarias locales y los órganos equivalentes del gobierno, constituye la clave del establecimiento y desarrollo de un sistema de información moderno que permitiese la notificación obligatoria de enfermedades y fuese capaz de generar información, analizarla, diseñando y aplicando intervenciones en materia de salud pública.

En 1925 apareció el *Boletín Técnico de la Dirección General de Sanidad* bimestral por medio de la Real Orden del 25 de Noviembre, con un formato similar al de boletines anteriores. Este Boletín Técnico de la Dirección General de Sanidad es el antecedente de la *Revista de Sanidad e Higiene Pública* (en 1987), que posteriormente pasó a tener el nombre actual de *Revista Española de Salud Pública* (en 1995).

Segundo paradigma de la Bioestadística en España: Marcelino Pascua Martínez

Marcelino Pascua Martínez (1897–1977) nació en Valladolid, ciudad donde comenzó los estudios de Medicina que finalizó en Madrid en 1925. En 1929 se hizo cargo de la *Sección de Estadísticas Sanitarias* del Boletín Técnico de la Dirección General de Sanidad.

En 1930, a iniciativa de Marcelino Pascua Martínez (1897–1977), se unifica la forma de obtener las tasas de mortalidad para establecer comparaciones entre las diversas capitales y provincias. La política científica de la Junta para Ampliación de Estudios e Investigaciones Científicas aplicada desde 1907 y el Patrocinio de la Fundación Rockefeller permitió que Marcelino Pascua desarrollara sus investigaciones entre 1927 y 1929 en la Universidad Johns Hopkins de Baltimore y en la University College y Nacional Institute of Health de Londres. Siendo discípulo de Pearl y Frost en Baltimore y de Greenwood en Londres, lo que le permitió tomar contacto con los precursores de la aplicación de los métodos estadísticos en la Epidemiología.

Los conocimientos estadístico-sanitarios recibidos posibilitan que el gobierno republicano español le nombre responsable de la Dirección General de Sanidad desde 1931 hasta 1933, con el objetivo de reformar la Administración sanitaria incorporando los métodos cuantitativos en la obtención de información de salud pública[35].

La reforma sanitaria de Marcelino Pascua se inició con la creación de unidades dependientes de la Dirección General de Sanidad. Una de las que más relevancia tuvo por la intensa actividad investigadora fue la *Comisión Permanente de Investigaciones Sanitarias*, creada en virtud del Real Decreto de 13 de octubre de 1931. A través de este organismo se

[35] Sus ideas reformistas las expuso en una ponencia ante el Congreso Nacional de Sanidad celebrado en 1934 y en la Real Orden de 15 de febrero de 1936.

desarrollaban áreas de investigación que las circunstancias epidemiológicas y sanitarias del país requiriesen y se potenciaban las relaciones de cooperación entre profesionales sanitarios españoles y latinoamericanos. Otras unidades importantes fueron la Sección de Higiene infantil y la de Higiene Social y Propaganda. En enero de 1933 el doctor Pascua creó la Secretaría General Técnica, atribuyéndole entre otras competencias la coordinación de las Secciones de la Dirección General de Sanidad, la traducción de textos científicos extranjeros, la formación de una biblioteca especializada, etc. Uno de los elementos necesarios en la Sanidad Pública española era la implantación de un modelo de asistencia médica colectivizada mediante la creación de un seguro de enfermedad aplicable a grandes masas de la población, no obstante este objetivo no fue conseguido por el doctor Pascua Martínez.

La tabla de vida elaborada por Merino fue reemplazada en 1942 por la que confeccionó Marcelino Pascua[36] aplicando la técnica diseñada por Reed y Merrel[37]. Con la publicación de este trabajo, Marcelino Pascua completa la investigación sobre demografía española que había iniciado en 1934 que comprendía diferentes estudios estadísticos: Mortalidad específica en España. I – Cálculo de poblaciones (1934) y Mortalidad específica en España. II – Mortalidad por sexos, grupos de edad y causas, en el periodo 1911–1930.

De su labor investigadora destacan de modo especial las monografías referidas al análisis de series temporales de mortalidad y morbilidad en diversos países europeos, para el período 1900–1950, que fueron publicadas por el doctor Pascua entre 1949 y 1951[38]. El establecer la intensidad habitual o esperable como patrón de la evolución de la enfermedad sin tener en cuenta el número de casos es, a juicio de Marcelino Pascua, uno de los temas más polémicos de la epidemiología, lo que le lleva a revisar el concepto tradicional de epidemia.

En 1953 ocupó el cargo de Director Consultor con la función de asesorar a los Gobiernos acerca de los mecanismos de mejora de los servicios nacionales de estadística sanitaria[39].

Durante el período de la guerra civil se desplazó exiliado a Suiza, Chile y Estados Unidos, volviendo posteriormente a Suiza para trabajar en el Departamento de Estadísticas Sanitarias de la Organización Mundial de la Salud desde 1948 hasta su jubilación en 1957.

Una vez jubilado elabora un Manual de Bioestadística que será publicado en 1965[40], que contenía siete capítulos, dos apéndices, un anexo con tablas de la distribución normal, t de Student y F de Snedecor. En este libro define el concepto de Bioestadística, recomienda la utilización en las investigaciones de datos estadísticos originales; indica las diferentes formas de presentar la información estadística, números índices y tasas aplicables en Medicina (natalidad, mortalidad, mortalidad proporcional, mortalidad infantil, morti-natalidad, mortalidad en la niñez, morbilidad y ataque secundario, letalidad por casos, de incremento natural, nupcialidad, tasas estandarizadas, tasas de mortalidad media equivalente, supervivencia y expectativa de vida, tasa de mortalidad potencial, método de observación continuada o progresiva por tasa de supervivencia…) Medidas de tendencia central, de dispersión, de variación. Series temporales, análisis de la varianza y el resto de materias que habitualmente aparecen en los textos que se utilizan actualmente.

---

[36] PASCUA MARTÍNEZ, M. (1942) *Mortalidad específica en España. III - Mortalidad por sexos y causas de defunción de la lista larga internacional en el período 1901–1930. IV - Tablas de vida*. The Johns Hopkins University, Baltimore.
[37] REED, L.J.; MERREL, M. (1939) “A short method for constructing and abridged life-table”. *American Journal of Hygiene*, 30 (2)
[38] PASCUA, M. (1950) “Evolution de la mortalité en Europe pendant le vingtième siècle”. *Rapport Epidémiologie Démographique*, 3, págs. 30 – 62.
[39] Noticias. *Crónica de la Organización Mundial de la Salud (1957)*, 11, pág. 181.
[40] PASCUA, M. (1965) *Metodología bioestadística para médicos y oficiales sanitarios*. Paz Montalvo, Madrid.

Capítulo 29

La regresión por mínimos cuadrados parciales: orígenes y evolución

GREGORIA MATEOS-APARICIO MORALES
Universidad Complutense de Madrid
ANTONIO JESÚS CABALLERO DOMÍNGUEZ
Jefe de Investigación de CFI Group de Madrid

Introducción

La técnica de regresión PLS o PLSR, Regresión por mínimos cuadrados parciales (Partial Least Squares Regression), se desarrolla para evitar, entre otros, el efecto de la multicolinealidad en las estimaciones de los parámetros de una regresión. El problema de la multicolinealidad, o relación de dependencia extrema entre variables explicativas en un modelo de regresión, produce situaciones de inestabilidad de los coeficientes de regresión. Los coeficientes de regresión pueden ser no significativos cuando las variables explicativas están muy correlacionadas con la variable explicada, produciendo dificultades de interpretación de la ecuación de regresión a causa de signos erráticos en los coeficientes.

Este problema suele surgir cuando se usa una cantidad relativamente grande de variables explicativas, y su tratamiento más directo es reducir la dimensionalidad de dicho conjunto de variables. La pregunta que inmediatamente surge es cómo hacer dicha reducción. La respuesta suele ser buscar una serie de nuevas variables creadas como combinación lineal de las originales, de forma que se elimine el problema de la multicolinealidad. El método de Componentes Principales ha sido muy usado durante años, y hasta hace poco era el referente entre las técnicas de reducción de la dimensionalidad. Este enfoque del método de Componentes Principales aplicado a la Regresión solía denominarse Regresión por Componentes Principales o PCR (Principal Components Regression).

Por su parte, el objetivo perseguido por la modelización PLS es la predicción de las variables dependientes. Este objetivo se traduce en un intento por maximizar la varianza explicada de dichas variables, por lo que PLS puede ser más adecuado para fines predictivos (Chin et al., 2003). En efecto, Wold (1979) afirma que PLS se orienta principalmente para el análisis causal predictivo en situaciones de alta complejidad pero con un conocimiento teórico poco desarrollado. También Barclay et al. (1995) concluyen que PLS se recomienda generalmente en modelos de investigación predictivos.

Las dos técnicas se comparan para la solución del problema de multicolinealidad en regresión logística funcional: uno basado en la regresión por componentes principales y otro basado en regresión PLS. PLS y PCR pretenden reducir la dimensionalidad, y, por tanto, hacer frente al problema que en muchas ocasiones presenta un conjunto de variables explicativas con un alto grado de multicolinealidad. Sin embargo, el enfoque de ambas técnicas es distinto, y por tanto, los resultados obtenidos también. Es decir, mientras que el propósito de PCR es recoger la máxima variabilidad o varianza de las variables explicativas, PLS trata de realizar lo mismo pero teniendo en cuenta además la relación entre X e Y. En otras palabras, PLS es un método más orientado a la predicción que PCR, ya que este último se centra en la reducción de la dimensionalidad de X (conjunto de variables explicativas) sin tener en cuenta la relación existente entre X e Y.

Algunos autores como Jerome Friedman o Trevor Hastie aseguran que análisis más profundos revelan que en el aspecto de recoger la máxima varianza en el problema de optimización, PLS tiende a dominar, y que, por tanto, PLS se comporta más parecido a PCR de lo que podríamos imaginar en un primer momento. Sin embargo, y a pesar de ello, en condiciones generales PLS aventaja a PCR como método predictivo de reducción de la dimensionalidad.

Orígenes y evolución de la regresión PLS

El método PLS fue desarrollado por el profesor sueco Herman Wold (mentor de Karl Jöreskog, fundador del Structural Equation Modelling). En 1966 Herman Wold publicó un libro en el que presentaba formalmente el método PLS, (Wold, H., 1966), aunque entonces se denominó NIPALS (Nonlinear Iterative Partial Least Squares) (Wold, 1973), y posteriormente PLS (Wold, 1979; 1982, 1985). El diseño básico de esta técnica terminó de completarse en 1977 (Wold, 1982), y éste se ha ido ampliando en siguientes etapas (Chin, 1998b) en las que continuó aplicándose a nuevos problemas o campos.

El método PLS ha tenido una difícil evolución, ya que al principio costó mucho incluirlo en un contexto estadístico y ello ralentizó su aplicación. Herman Wold presentó por primera vez su metodología PLS aplicada a los modelos de rutas con variables latentes en 1979, aunque fue varios años más tarde, en 1982 y 1985, cuando publicó sus principales referencias de algoritmos.

Svante Wold, hijo de Herman Wold, y Harold Martens adaptaron NIPALS en 1983 para resolver el problema de la multicolinealidad en las estimaciones de los parámetros de los modelos de regresión lineal. A esta adaptación la llamaron Regresión PLS. Svante Wold, continuó la labor de su padre, aunque más centrado en el campo de la Quimiometría.

En 1990, Stone y Brooks situaron el método PLS en el campo de la "Regresión Continua", y significó el primer paso serio dado por estadísticos para situar PLS en un entorno estadístico.

El siguiente paso fue el dado por Ildiko Frank y Jerome Friedman al presentar este método como una herramienta importante en Química y Quimiometría. Fue en 1993 cuando estos dos autores lograron definitivamente situar al método PLS dentro de un contexto estadístico, más concretamente en el área del Análisis Multivariante clásico. Esta contextualización fue criticada por Svante Wold, que apoyó la publicación del trabajo de Hinkle y Rayens en el que se referían al método PLS como una técnica “composicional” de datos. Esto produjo una importante discusión en todo el mundo sobre el tipo de técnica que PLS representaba, pero esto no entorpeció su desarrollo sino que logró impulsarlo. Así, apareció una versión no lineal de PLS, dando lugar a la aparición de una innumerable cantidad de algoritmos PLS aplicados a distintos problemas. En la actualidad, PLS se ha convertido en una técnica que por sí sola reúne suficiente conocimiento como para ser un campo en sí mismo, y nuevos algoritmos van apareciendo día a día, dando lugar a técnicas como la Regresión PLS Logística.

La técnica PLS además de ser una solución al problema de la multicolinealidad en un modelo de regresión, resuelve otro problema que es el que se presenta cuando el número de individuos es menor que el número de variables y cómo ello afecta a la estimación de los coeficientes de regresión. Svante Wold fue capaz de realizar mediante PLS una regresión con 10 variables y un único caso, es decir con una muestra de tamaño uno, frente a los cinco por variable que nos recomienda la estadística tradicional. Esto nos da idea de la potencialidad del método en situaciones con muestras pequeñas. PLS puede ser una poderosa herramienta de análisis por las mínimas exigencias en términos de escalas de medición, tamaño de las muestras y distribuciones residuales. PLS no precisa que los datos provengan de distribuciones normales o conocidas (Falk y Miller, 1992). La metodología PLS no impone grandes restricciones al modelo, y ello redunda en una clara simplificación de la teoría necesaria para su manejo, y del tiempo de aprendizaje. Pero hay que advertir que aunque la hipótesis de normalidad de los datos raramente se encuentra en la realidad, y a pesar de que con esta técnica se puede obviar esta restricción, los resultados y decisiones basadas en ellos quedan claramente comprometidos.

Gracias al avance informático de los últimos años, ha sido posible plantear técnicas de remuestreo o computación intensiva que nos permitan validar las estimaciones obtenidas en nuestros modelos. Tres de estas técnicas comúnmente utilizadas en modelos PLS son: la validación cruzada (o *blindfolding* en inglés), la de Jacknife y la de Bootstrap. El método Bootstrap fue desarrollado por Efron en 1979 y requiere de métodos de simulación para la estimación de un parámetro y de su varianza. Aunque el rendimiento del método de Bootstrap es mejor, en general, que el de Jacknife, este último es más eficiente en términos de tiempo de computación. Chin es de la opinión que en modelos PLS, las estimaciones dadas por uno o por otro deben converger.

Distintos enfoques de la Regresión PLS

Existe gran confusión en lo referente a PLS en la actualidad, tanto de autoría (en cuestiones de asignar la propuesta como mérito del padre o del hijo), como de enfoque, Regresión PLS o PLS- Path Modeling. Por ello, revisamos estos enfoques para poder entender su relación, como parte de la evolución de esta técnica.

Wold estaba en desacuerdo con la aproximación dura o *hard modeling* que Jöreskog propugnaba para los modelos con rutas y variables latentes. El enfoque de Jöreskog necesitaba de la imposición de fuertes hipótesis sobre la distribución de los datos, y de un número elevado de casos o tamaño muestral. Sin embargo, el enfoque de Wold era mucho más ligero, y a la práctica ausencia de hipótesis sobre la distribución de los datos se unía la ventaja de que un reducido número de casos podía ser suficiente para usar el algoritmo. Esta era la razón de que al enfoque de Wold se le denominara *Soft Modeling* o aproximación blanda.

Ambos enfoques sobre los modelos de rutas o modelos de ecuaciones estructurales fueron comparados por ambos en 1982 en "*Soft Modeling: The basic Design and Some Extensions*" en Systems under Indirect Observation – Causality Structure Prediction. La gran conclusión de dicha comparación es que, realizando algunos cambios en el algoritmo LISREL de

Jöreskog, las estimaciones dadas por ambos enfoques están correladas, aunque con el algoritmo original se conseguían estimaciones más robustas que con el método PLS.

Años más tarde de esta publicación, y en vista de la gran difusión que tuvo el enfoque PLS, Harald Martens propuso llamar PLS-Path Modeling al enfoque PLS sobre modelos de ecuaciones estructurales.

Uno de los problemas que surgen en modelos de ecuaciones estructurales (SEM) es su indeterminación. Es decir, existe un número de parámetros a estimar demasiado grande para el tamaño muestral de trabajo. Esto es debido principalmente a que las variables latentes son totalmente desconocidas e introducen una fuerte deficiencia en el modelo a la hora de medir las relaciones existentes entre ellas. Mediante PLS, este problema es resuelto de una forma muy sencilla, creando simplemente las variables latentes como sumas ponderadas de las variables manifiestas correspondientes. Esto permite dos cosas; una: resolver, como decíamos, la indeterminación del modelo al tener las variables latentes, y otra: poder analizar las puntuaciones obtenidas en dichas variables latentes. Esto último resulta muy interesante para las compañías, ya que cada variable latente suele reflejar un proceso de percepción de ellas, y poder compararla con las de la competencia es siempre deseable.

La metodología PLS-Path Modeling presupone que los modelos estructurales son lineales, por lo que cualquier técnica de regresión es susceptible de ser usada para la estimación de los coeficientes estructurales. Sin embargo, el método de regresión por mínimos cuadrados ordinarios es el más extendido debido a la falta de requisitos para su aplicación. De esta forma, la medida usada de bondad del ajuste es la propia de este tipo de modelos: el coeficiente de determinación R^2, esto es, el cociente entre la variabilidad explicada por la regresión y la variabilidad total.

Referencias de Algoritmos de la Regresión PLS

En primer lugar hay que diferenciar el caso en el que la variable explicativa Y es univariante y el caso de cuando es multivariante, ya que existen dos métodos para la Regresión PLS, PLS1 y PLS2, adaptados a estas circunstancias. En la Regresión PLS1 hay tan sólo una variable a explicar y p explicativas, mientras que en la Regresión PLS2 hay q variables a explicar ($q>1$) y p variables explicativas.

Existen muchos algoritmos de regresión PLS en la actualidad, pero sin duda alguna, el más útil para PLS-Path Modeling es el algoritmo NIPALS. Este algoritmo fue creado, como ya hemos referido, por Herman Wold en 1966 y ofrece la gran ventaja de no necesitar ni suprimir ni estimar los datos que sobran o faltan de una observación para que se pueda usar en el análisis. Su denominación se debe a que permite estimar los parámetros de un modelo no lineal mediante una serie de regresiones simples entre los datos y una parte de los parámetros, de ahí su nombre NIPALS (Nonlinear estimation by Iterative Partial Least Squares) que en español quiere decir algoritmo no lineal de estimación por mínimos cuadrados parciales iterados.

Sin duda alguna, el autor que de una forma más clara expone el algoritmo para esta técnica es Claes Fornell en su obra "*A Second generation of multivariate Análisis Vol. I*". A diferencia de otros autores como Löhmoller, Wold o Tenenhaus, Fornell se centra en la aplicabilidad del método en lugar de en la teoría matemática subyacente a dicho método. Para la mayoría de los investigadores este enfoque es el más útil, puesto que tanto la notación como el desarrollo matemático de otros enfoques es bastante engorroso y poco aclaratorio sobre el procedimiento que sigue el método. Por ello, tomaremos como referencia el algoritmo ofrecido por Fornell en la obra mencionada para ilustrar el algoritmo PLS-Path Modeling.

Otro algoritmo que cabe destacar es el elaborado por Claes Fornell para estudios de Satisfacción. Dicho algoritmo convierte las variables (siempre medidas en una escala 1-10) a una escala 0-100. Fornell ha logrado con este enfoque medir la economía estadounidense en términos de Satisfacción. De hecho, su metodología es la que se sigue en la actualidad en la medición de activos intangibles.

Una de las grandes ventajas que ofrece para el investigador la técnica PLS-Path Modeling es la simplicidad en la lectura y diagnosis del modelo (en comparación con otras técnicas). Esta simplicidad no excluye la modelización necesaria para cualquier estudio estadístico, pero sí que ayuda a la hora de reducir el tiempo de aprendizaje para el investigador, puesto que se detiene más que en consideraciones técnicas o algorítmicas en su aplicación a problemas o situaciones.

Evolución del software para el modelo PLS

El Dr. H. Martens, quién diseñó el primer paquete de software para PLS a mediados de los ochenta, propuso la idea y la aplicación de una forma avanzada de PLS que se designó L-PLS. Con el uso de PLS se pueden conectar entre sí datos de distinta índole: atributos de un producto, datos de preferencias y datos demográficos de los consumidores. Si se combinan estos grupos de datos, podemos obtener un mapa de preferencias que describa como se interrelacionan los atributos, los datos de preferencias y los datos demográficos de los consumidores: como edad, estatus económico o género. De este modo, podemos averiguar el grupo de consumidores que aceptaría un nuevo producto y modificar o diseñar un producto que satisfaga los gustos de un grupo de consumidores objetivo. Por supuesto el L-PLS también puede analizar otras combinaciones.

La disponibilidad de software para llevar a cabo un análisis PLS es limitada. Durante años el único software disponible para esta modelización flexible de la técnica PLS, PLS-Path Modeling, fue LVPLS 1.8 elaborado por Jan Bernd Löhmoller (1984). El propio Löhmoller tuvo su gran aportación en “Latent Variable Path Modeling with Partial Least Squares” en 1989, donde amplió las bases del algoritmo PLS en varias direcciones.

Más recientemente han aparecido nuevos programas comerciales, en versión para Windows, como PLS-Graph de Wynn Chin (Chin, 2004), profesor de la Universidad de Houston, con un interfaz intuitivo y amigable, siendo el profesor Wynne Chin la persona que está liderando este proyecto. PLS-Graph ha sido ya empleado en diversos trabajos de investigación.

También, en versión para Windows, existen aplicaciones gratuitas como SmartPLS o Visual PLS (versión gráfica del programa de Löhmoller), un interfaz gráfico muy sencillo de usar y muy intuitivo, totalmente disponibles en Internet. Este último diseñado por Jen Ruei Fu, profesor de la Universidad taiwanesa de Kaohsiung. La principal ventaja de esta herramienta es que es totalmente gratuita y su fiabilidad ha sido bien contrastada. Visual-PLS puede ser descargado de la página. http://www2.kuas.edu.tw/prof/fred/vpls/

La limitada difusión de los programas para PLS, junto con el carácter difícil y poco amigable que posee LVPLS, ha sido lo que ha llevado a lo no muy amplia difusión de la metodología PLS con relación a la otra rama de los modelos de ecuaciones estructurales (SEM). Sin embargo, creemos que cada vez más se está ampliando la base de investigadores que están aplicando esta técnica, a lo cual contribuirá la distribución de PLS-Graph. Para más información se puede consultar la dirección de Internet: http://disc-nt.cba.uh.edu/chin/indx.html.

Finalmente, podemos avanzar el desarrollo de otros programas que permiten implementar este tipo de modelización flexible. Prueba de ello, son las últimas demostraciones realizadas en el 3rd International Symposium on PLS and Related Methods (Lisboa, 2003): http://www.isegi.unl.pt/pls03/default.htm

La técnica PLS en las Ciencias Sociales

El enfoque del Modelo de Ecuaciones Estructurales o SEM (Structural Equation Modeling) basado en las covarianzas (especialmente bajo la estimación de máxima verosimilitud) busca encontrar parámetros invariantes estructurales o funcionales que definan cómo funciona el mundo, es decir, persiguen proporcionar una afirmación de causalidad, una descripción de los mecanismos causales (modelización firme o rígida). El problema que surge al intentar alcanzar este conocimiento con estas técnicas son las hipótesis restrictivas que se requieren. Las distribuciones de los datos y los niveles de medida de las variables se pueden encontrar dentro de lo que se define como un sistema cerrado (Falk y Miller, 1992).

Sin embargo, dados estos requerimientos, parece difícil la aplicación estricta de este tipo de modelización en el campo de las ciencias sociales. En esta situación surge PLS, técnica que tal y como fue diseñada refleja las condiciones teóricas y empíricas de las ciencias sociales y del comportamiento, donde son habituales las situaciones con hipótesis teóricas no suficientemente asentadas y escasa información disponible (Wold, 1979).

PLS representa un tipo de sistema de análisis matemático y estadístico de datos que se adecua a las condiciones presentes en las ciencias sociales. En esta forma de modelización, conocida como modelización flexible (Wold, 1980) los procedimientos matemáticos y estadísticos subyacentes en el sistema son rigurosos y robustos (Wold, 1979) y, sin embargo, el modelo matemático es flexible en el sentido de que no realiza suposiciones relativas a niveles de medida, distribuciones de los datos y tamaño muestral. El objetivo que se persigue es más moderado que la modelización firme.

En la modelización flexible se crean relaciones predictivas lineales óptimas entre variables. “En un sentido de mínimos cuadrados, esto significa que dados los datos y el modelo, las variables independientes se vuelven las mejores variables predictoras posibles, y las variables dependientes se vuelven las mejores variables criterio o predecidas” (Falk y Miller, 1992, p. xi). Como se observa, se abandona la idea de causalidad (presente en la modelización firme) y se reemplaza por el concepto de predictibilidad. Mientras que la causalidad garantiza la capacidad de controlar los acontecimientos, la predictibilidad permite sólo un limitado grado de control (Falk y Miller, 1992).

Podemos constatar que la modelización flexible es un método para estimar la probabilidad de un acontecimiento en función de la información disponible sobre otros acontecimientos. No pretende ser un sistema de valoración de la causalidad, pero es particularmente aplicable cuando no se producen las condiciones de un sistema cerrado.

Bagozzi fue uno de los primeros autores que planteaba el uso de técnicas PLS para los modelos de Satisfacción desde un punto de vista de negocio o marketing. Tras él, vinieron otros como Claes Fornell, referido anteriormente, o Anders Westlund que desarrollaron una metodología propia para los estudios de Satisfacción y que incluía desde la fase cualitativa (en la que se obtienen los atributos o variables que conforman el cuestionario) hasta la interpretación de los resultados. Los estudios de Satisfacción suponen actualmente un gran marco de referencia en la Investigación de Mercados de la empresa moderna, y en su estrategia de relación con el cliente.

Sin embargo, la solución que ofrece PLS a la indeterminación del modelo tiene un precio; los estimadores de los parámetros no son consistentes y aparece un sesgo en sus estimaciones. Recordar que el estimador de un parámetro es consistente si se aleja del parámetro con una probabilidad débil, si el tamaño de la muestra es lo suficientemente grande. En realidad el precio que tenemos que pagar es relativamente pequeño, ya que si bien no se logra en principio la consistencia, sí está probada la consistencia cuando el tamaño muestral es suficientemente grande, es decir, la consistencia es asintótica. En los modelos SEM, la consistencia está asegurada si las hipótesis que se imponen se cumplen, lo cual es realmente extraño en la realidad. En cualquier caso, la consistencia es un propiedad asintótica tanto en modelos SEM, como en modelos PLS o de otro tipo, por lo que es sólo una cuestión de tamaño muestral. Wold siempre defendió el uso de técnicas no paramétricas para la evaluación de la consistencia en los modelos PLS.

En resumen, las mínimas exigencias en la distribución de la población y el tamaño muestral del modelo PLS han ocasionado que éste se popularice mucho entre los investigadores de administración de empresas, sobre todo por las muchas ventajas que presenta frente las técnicas basadas en la covarianza.

Bibliografia

BARCLAY, D.; HIGGINS, C.; THOMPSON, R. (1995): “The Partial Least Squares (PLS) Approach to Causal Modelling: Personal Computer Adoption and Use as an Illustration”, Technology Studies, Special Issue on Research Methodology, 2(2): 285-309.

CABALLERO DOMÍNGUEZ, ANTONIO JESÚS. “SEM vs. PLS: Un enfoque basado en la práctica”, IV Congreso de Metodología de Encuestas. Universidad Pública de Navarra. Septiembre 2006

CHIN, W.W. (1998b): “The Partial Least Squares Approach to Structural Equation Modelling”, en G.A. Marcoulides [ed.]: *Modern Methods for Business Research*, pp. 295-336. Mahwah, NJ: Lawrence Erlbaum Associates, Publisher.

CHIN, W.W.& NEWSTED, P.R. (1999): “Structural Equation Modelling Analysis with Small Partial Least Squares”, en R. Hoyle [ed.]: *Statistical Strategies for Small*

CHIN, W. W., (2000). Frequently Asked Questions – Partial Least Squares & PLSGraph. Home Page.[On-line]. Available: http://disc-nt.cba.uh.edu/chin/plsfaq.htm

CHIN, W.W.; (2004): *PLS-Graph. Version 3.00*. build 1060. University of Houston, Texas, USA.

EDWARDS, J. R., & BAGOZZI, R. P. (2000): “On the nature and direction of relationships between constructs and measures”, Psychological Methods, 5(2): 155-177.

EFRON, B. (1982): *The Jackknife, The Bootstrap and Other Resampling Plans*. Philadelphia: Society for Industrial and Applied Mathematics.

EFRON, B. & GONG, G. (1983): “A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation”, The American Statistician, 37: 36-48.

FALK, R.F. & MILLER; N.B. (1992): *A Primer for Soft Modelling*. Akron, Ohio: The University of Akron.

FORNELL, C. (1982): "A Second Generation of Multivariate Analysis: An Overview", en C. Fornell [ed.]: *A Second Generation of Multivariate Analysis, 1*: 1-21. New York: Praeger Publishers.

FORNELL, C. & BOOKSTEIN, F.L. (1982): "A Comparative Analysis of Two Structural Equation Models: Lisrel and PLS Applied to Market Data", en C. Fornell [ed.]: *A Second Generation of Multivariate Analysis, 1*: 289-324. New York: Praeger Publishers.

FRANK, I.E., & FRIEDMAN, J.H. (1993): "A statistical view of chemometrics regression tools". Technometrics, 35, 109–148.

JÖRESKOG, K.G. & WOLD, H. (1982): *Systems under Indirect Observation – Causality Structure Prediction*. Amsterdam: North Holland Publishing Company.

JÖRESKOG KARL G. (1993). *Modelado de ecuaciones estructurales con LISREL*. Instituto Vasco de Estadística.

JÖRESKOG, KARL G & SÖRBOM, DAG. (1993). *LISREL 8. Structural Equation Modelling with the Simplis Command Language*. Lawrence Erlbaum Associates Publishers.

JÖRESKOG, KARL G & SÖRBOM, DAG. 1993. LISREL 8 y PRELIS 2. *User´s Reference Guide*. SSI. Scientific Software International.

LOHMÖLLER, J.B. (1984): *LVPLS Program Manual. Version 1.6. Latent Variables Path Analysis with Partial Least-Squares Estimation*. Köln: Zentralarchiv für Empirische Sozialforschung, Universität zu Köln.

WOLD, H. (1966): "Estimation of principal components and related models by iterative least squares". In P.R. Krishnaiaah (Ed.). *Multivariate Analysis*. (pp.391-420) New York: Academic Press.

WOLD, H. (1973): "Nonlinear Iterative Partial Least Squares (NIPALS) Modelling: Some Current Developments", en P.R. Krishnaiah [ed.]: *Multivariate Analysis: II*, Proceedings of an International Symposium on Multivariate Analysis Held at Wright State University, Dayton, Ohio, June 19-24, 1972, pp. 383-407. New York: Academic Press-

WOLD, H. (1979): "Model Construction and Evaluation when Theoretical Knowledge Is Scarce: An Example of the Use of Partial Least Squares". Cahiers du Département D´Économétrie. Genève: Faculté des Sciences Économiques et Sociales, Université de Genève.

WOLD, H. (1980): "Soft Modelling: Intermediate Between Traditional Model Building and Data Analysis", Mathematical Statistics, 6: 333-346.

WOLD, H. (1982): "Systems Under Indirect Observation Using PLS", en C. Fornell [ed.]: *A Second Generation of Multivariate Analysis, 1*: 325-347. New York: Praeger Publishers.

WOLD, H. (1985): "Systems Analysis by Partial Least Squares", en P. Nijkamp, H. Leitner y N. Wrigley [ed.]: *Measuring the Unmeasurable*, pp. 221-251. Dordrecht: Martines Nijhoff Publishers.

WOLD H. (1985): "Partial Least Squares", in S. Kotz and N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences (vol. 6)*, New York: Wiley, pp. 581-591.

WOLD, S. (2001): "Personal memories of the early PLS development". Chemometrics and Intelligent Laboratory Systems, 58, 83–84.

Capítulo 30

Logit model.
De Verhulst (1838) a Mcfadden (2001)

ELENA MARTINEZ RODRIGUEZ
Universidad Complutense de Madrid

Introducción

Definimos la ecuación ó función logística mediante la expresión:

$$P(Y) = \frac{\exp^{Y}}{1+\exp^{Y}}, \tag{1}$$

en la que habitualmente P denota una función de probabilidad, e Y indica una combinación lineal del tipo $Y = \beta_1 + \beta_2 X_{2i} + ... + \beta_k X_{ki}$

Así definida, la función logística cumple las propiedades de ser una función monótona creciente, acotada en el intervalo $[0,1]$. Su representación gráfica (figura 1) es una curva de forma sinusoidal

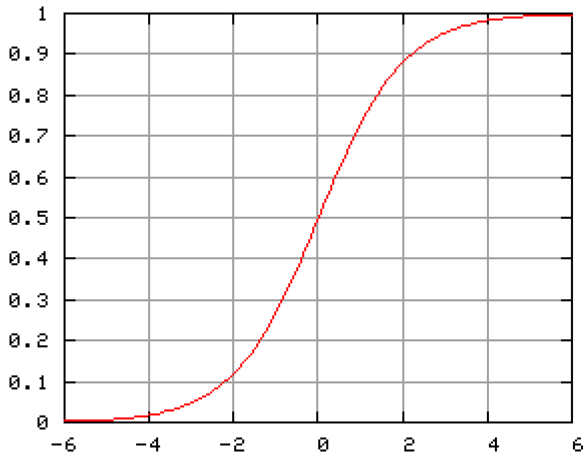


Figura 1: Función Logística

en la que observamos que para valores pequeños de la variable Y la función experimenta un crecimiento lento, que aumenta rápidamente a medida que aumenta la variable (en este tramo se asemeja al crecimiento exponencial), y, finalmente, se ralentiza para valores altos de *Y*, alcanzado su cota máxima situada en el valor 1.

El significado de esta ecuación depende de la definición de las variables. Por ejemplo, en Biología es frecuente que *P* sea la función de probabilidad del suceso dicotómico supervivencia o muerte del organismo observado cuando es sometido a un estímulo continuo, expresado éste mediante un modelo lineal (*Y*)

Siglo XIX: La Ecuación de Verhulst

Según Thompson, el término "curva logística" debe atribuirse a Edward Wright (1558-1615), quien usó este término para referirse a una curva o ecuación logarítmica.

No obstante debemos esperar hasta el siglo XIX para que la función logit o ecuación logística se desarrollara tal y como la conocemos hoy en día. Esta función se aplicó de forma independiente en dos ámbitos bien diferenciados: en química, para explicar reacciones autocatalíticas y en demografía, para explicar el crecimiento de poblaciones. Hay que destacar que la trayectoria en ambas disciplinas fue distinta: su uso en química ha sido ininterrumpido, mientras que en demografía hay un periodo prolongado en el que cae en un injustificado olvido.

Una reacción química es clasificada como autocatalítica o en cadena cuando el producto en sí mismo es el catalizador de la reacción. Son reacciones que se caracterizan porque al principio proceden lentamente debido a la escasa presencia del elemento catalizador, aunque va aumentando progresivamente para, al final, sufrir un retroceso al disminuir la concentración del reactivo. La función logística, tal como hemos visto en el apartado anterior, describe la evolución temporal de este tipo de reacciones. Prueba de ello son los trabajos, fechados en 1883, del químico alemán W. Ostwald.

El otro campo en el que se utilizó la ecuación logística para describir un fenómeno fue la demografía. En concreto, la ecuación logística de crecimiento de una población fue propuesta por P. Verhulst como una posible solución al dilema del crecimiento exponencial de Malthus.

A finales del siglo XVIII el economista inglés Thomas Robert Malthus, publicó la obra *Ensayo sobre el principio de la población* (1798), en la que expuso y defendió sus teorías sobre el crecimiento demográfico, según las cuales la población humana tiende a crecer en progresión geométrica mientras que los medios de subsistencia lo hacen en progresión aritmética, principio que hoy conocemos como "Ley exponencial de crecimiento poblacional". Este modelo podemos representarlo por la siguiente expresión:

$$W(t) = W(0)\, e^{\beta t}, \qquad (2)$$

siendo $W(t)$ el tamaño de la población en el instante t, $W(0)$ el tamaño inicial y β la tasa constante de crecimiento.

Alfhonse Quetelet, matemático y astrónomo belga (1795-1874), fue uno de los primeros en considerar que el modelo exponencial de crecimiento no era adecuado para explicar la expansión demográfica de un país. Si bien podía reflejar la situación de Estados Unidos a principios del siglo XIX (modelo poblacional observado por Malthus), un país joven y casi vacío, su aplicación a otras sociedades conduciría a valores imposibles. Quetelet estaba convencido de que una población no podía crecer indefinidamente, sino que existían fuerzas, tanto externas como internas, que tienden a prevenir este crecimiento. Aunque sus aportaciones en dinámica de la población no son destacables, la referencia a este autor es obligada porque fue profesor, en la Universidad de Gante, de Pierre-François Verhulst (1804-1849) con quien trabajó durante un largo periodo y sobre quien tuvo una gran influencia tanto en su vida personal como en su obra. Quetelet fue pionero en la aplicación de la Estadística y la Teoría de la Probabilidad para explicar las regularidades sociales (concepto del “hombre medio”). Su preocupación por el orden social y el papel que le atribuía a la ciencia como un instrumento fundamental para su control, son determinantes en la formación y en la trayectoria académica y científica de su discípulo.

Verhulst abordó el problema del crecimiento de una población adoptando las hipótesis de Quetelet y, por tanto, consideró que es un proceso limitado. Demostró que la tasa de crecimiento de una población está limitada directamente por su propia densidad. Por este motivo añadió al modelo propuesto por Malthus un término adicional que representa la resistencia al crecimiento:

$$W(t) = \beta W(t) - \phi(W(t)) . \tag{3}$$

En concreto, $\phi(W(t))$ debía representar las fuerzas que frenan el crecimiento poblacional y que ambos autores consideraban que aumentan con el cuadrado de la tasa de variación de la población. Esta idea se puede expresar mediante la ecuación:

$$W(t) = \beta W(t)(\Omega - W(t)) , \tag{4}$$

donde Ω denota el límite máximo o nivel de saturación de la magnitud tamaño poblacional *W*. Observamos en (4) que el crecimiento es proporcional tanto al tamaño poblacional en el momento t, como al crecimiento potencial de la población $(\Omega - W(t))$.

Si expresamos *W (t)* como proporción, tendremos:

$$P(t) = \frac{W(t)}{\Omega} = \beta P(t)\left[1 - P(t)\right]. \tag{5}$$

La solución a esta ecuación diferencial no lineal la encontramos en la expresión (6), a la que Verhulst llamó función logística:

$$P(t) = \frac{\exp^{(\alpha+\beta t)}}{1+\exp^{(\alpha+\beta t)}} . \tag{6}$$

De esta forma, el tamaño de una población en el instante *t* se calcularía como:

$$W(t) = \Omega \frac{\exp^{(\alpha+\beta t)}}{1+\exp^{(\alpha+\beta t)}} , \tag{7}$$

ecuación que en Biología se denomina *ecuación de Verhulst*, en honor a su autor, y que describe un modelo de crecimiento autolimitado de una población, conocido también como modelo logístico de crecimiento poblacional.

Como veremos a continuación, Verhulst aplicó su ecuación para calcular el límite superior de la población de algunos países europeos, entre ellos Bélgica, situando el nivel de saturación en 9,5 millones de habitantes. La población belga en 1994 era de 10,118 millones de habitantes; si descontamos el efecto de la inmigración (no contemplado en el modelo (7)) parece que la predicción de Verhulst es bastante acertada.

Los resultados de sus investigaciones sobre el crecimiento demográfico vieron la luz a través de varios artículos. En 1838 publica en *Correspondance Mathématique et Physique* (editado por A. Quetelet) «*Notice sur la loi que la population suit dans son accroissement*», en el que expone la esencia de su teoría y muestra cómo su modelo describe fielmente el crecimiento de las poblaciones de Bélgica, Rusia y Francia, utilizando para ello datos anteriores a 1833. No obstante, en este artículo no menciona cómo deduce la ecuación de crecimiento ni tampoco se refiere a ella como función logística.

En 1845 publica, en *Nouveaux Mémoires de l'Académie Royale des Sciences de Belgique*, un segundo artículo en el que por primera vez introduce el término "logístico" para referirse a la ecuación de crecimiento poblacional (7), proporcionando más detalles sobre sus propiedades, incluso llega a representar la curva logística junto con la curva exponencial. A nivel aplicado, este artículo es el más completo ya que estima los parámetros α y β de la ecuación (7)[1], lo que le permite realizar una predicción para el tamaño máximo de la población (Ω) de Bélgica y de Francia. Para Bélgica estima ese límite máximo en 6,6 millones y para Francia en 40 millones.

El último artículo referido al modelo de crecimiento poblacional lo publica en 1849, también en *Nouveaux Mémoires de l'Académie Royale des Sciences de Belgique*. El objetivo de este artículo es publicar las notables correcciones en las estimaciones de los parámetros del modelo, que permiten una nueva predicción del nivel máximo de la población en el caso de Bélgica, situándolo en 9,5 millones.

Tras la muerte de Verhulst (1849), J. B. Liagre, amigo del matemático, repite, utilizando datos más actualizados, la estimación del límite máximo de crecimiento de la población de Bélgica, que recoge en la segunda edición de su libro *Calcul des probabilités et théorie des erreurs* (1879), llegando, en este caso, a la cifra de13,7 millones. Con la excepción de este trabajo, la obra de Verhulst permanece casi 70 años en el olvido.

Primera mitad del siglo XX: Redescubrimiento y desarrollo de la función Logit

La primera vez, a excepción del libro de Liagre, que se menciona la obra de Verhulst es en 1918, tras más de 60 años de silencio. En este año, el matemático francés Gustave Du Pasquier publicó un artículo en el que se limitaba a exponer distintas teorías matemáticas formuladas sobre el comportamiento poblacional. En concreto, las teorías de Halley, Moivre, Euler y Verhulst. El artículo no tuvo repercusión.

En 1920 el biólogo americano Raymond Pearl, en colaboración con el matemático L. J. Reed, redescubre la función logística como un simple modelo de crecimiento demográfico. Hablamos de redescubrimiento porque, a diferencia de Du Pasquier que tan sólo la menciona, Pearl la aplica en varios de sus trabajos.

Raymond Pearl es sin duda uno de los biólogos más importantes del siglo XX, a la vez que uno de los científicos más prolíferos de esta época, como atestiguan sus 712 artículos y más de 17 libros, escritos sobre temas tan diversos como genética, alcohol, tabaco, fertilidad, longevidad, alimentos y precios, estadística o crecimiento de la población, entre otros

[1] Para la estimación de los parámetros, tanto para el caso de la población belga como de la francesa, considera únicamente valores conocidos de la población en tres años distintos.

muchos. Después de su trabajo doctoral, Pearl comenzó a investigar sobre la aplicación de los métodos estadísticos en biología, en colaboración con Karl Pearson. Ambos se asociaron en 1906 para editar la revista *Biométrica*, asociación que duró hasta 1910. A partir de 1919 su interés se centra en temas sobre crecimiento poblacional, llegando a proyectar el crecimiento de la población de Estados Unidos hasta 1940. Para realizar esta proyección, Pearl y sus colaboradores se basan en datos de la población de este país registrados desde 1790 a 1910, y utilizan la ecuación (7) de crecimiento poblacional de Verhulst, en la que, al igual que el matemático belga, estiman los parámetros a partir de los datos poblacionales conocidos en tres periodos.

J. S. Cramer defiende, en el capítulo 9 de su libro *Logit Models. From Economiccs and Other Fields,* que Pearl y sus colaboradores llegaron de forma independiente a la formulación de la ecuación logit recogida en la expresión (7), y que utilizan en varias de sus publicaciones, ya que, aunque conocían la aplicación de la curva logística para explicar la reacciones químicas en cadena, en 1920 (primera fecha de publicación de una serie de artículos fundamentados en el modelo logístico de crecimiento) no conocían el modelo propuesto por Verhulst. En el siguiente trabajo publicado en 1922 aparece la primera y escueta referencia, como pie de página, al matemático Verhulst. Referencia que se repite, con algo más de detalle, en la publicación de 1923.

En 1925 encontramos dos aportaciones destacadas al desarrollo de la función logística. Por una parte, el estadístico G. U. Yule, que conoce la aportación de Verhulst gracias a los trabajos de Pearl, dedica una parte del artículo publicado en este año a comentar el modelo de crecimiento de Verhulst, en concreto, uno de los apéndices. Es también Yule quien reestablece el término de función logística, que no aparece en ninguno de los papeles de Pearl y Reed ni en el artículo de Du Pasquierer. Por otra parte, el biólogo Alfred J. Lotka obtuvo en el ámbito de la ecología la ecuación logística para explicar el crecimiento demográfico de una comunidad. Este autor la denomina “ley del crecimiento poblacional”.

El término “Logit Model” fue acuñado por Joseph Berkson en 1944. La formación inicial de este autor es en Ciencias Físicas, disciplina en la que se familiariza con la función logística, publicando, junto con Reed (1929), un artículo referente a las funciones que modelan las reacciones autocatalíticas. Posteriormente se interesó por la Estadística, en concreto, por la aplicación de los métodos estadísticos en el campo de la Biología (Biometría). Quizá una de sus aportaciones más conocidas sea el llamado *modelo del error de Berkson*, en el que considera, en contraposición al modelo clásico de regresión, que el error es independiente de la variable observada. Igualmente se le reconoce como un de los autores que más utilizaron la función logística, usando por primera vez la denominación de Logit Model, en el artículo *Application to the Logistic Function to Bio-assay*. En este artículo el autor explica que usa el término Logit para la expresión

$$\log\left(\frac{p_i}{1-p_i}\right) = \log it(P), \tag{8}$$

siguiendo a Ittner Bliss (1934), quien llamó Probit a una función análoga, aunque lineal en variable independiente, cuando asume como función de probabilidad la curva Normal en lugar de una función logística. Podemos considerar, por tanto, que es Berkson quien introduce el uso del Modelo Logit en la Estadística, aunque la generalización de su uso no fue ni fácil ni rápida.

Segunda mitad del siglo XX: Modelos de elección discreta de McFadden

La etapa de desarrollo del Modelo Logit y su aplicación en distintos campos no es sencilla. Existe, desde su introducción por Berkson, una "rivalidad" entre los partidarios de la utilización de Modelos Logit y de Modelos Probit, tal y como se puede comprobar en muchas publicaciones, sobre todo durante las décadas de los 60 y 70. A esta rivalidad contribuye el propio Berkson, mediante una serie de artículos que abarcan desde 1944 hasta 1980. Los partidarios del Modelo Logit argumentan la flexibilidad en su interpretación, ya que ésta depende de la definición de las variables, y su simplicidad de cálculo y aplicación, que se ve reforzada en los 80 con la revolución informática. Sus detractores, en cambio, utilizan esta falta de interpretación específica para designar al Modelo Logit como una simple herramienta estadística, llegando a afirmar que carece de base teórica sólida para su aplicación, ya que, a diferencia del Modelo Probit, no tiene asociada una distribución de probabilidad reconocida (Aitchis y Brown, 1957).

La utilización del Logit Model en distintas ciencias ocurre a un ritmo desigual e, incluso, de forma independiente. Los primeros avances de este modelo ocurrieron en Estadística y en Epidemiología.

En Epidemiología los estudios de casos y controles, basados en ratios de probabilidad, propician la rápida incorporación del modelo logit. De hecho, ya en los años 50, J. Cornfield utilizó la regresión logística para el cálculo de los *odds ratio* como valores aproximados del riesgo relativo.

En estadística el principal difusor de la regresión logística fue D. R. Cox con la publicación en 1970 de su libro *The Análisis of Binaria Data.* La importancia de los trabajos de este autor reside tanto en la utilización que hace del Modelo Logit para formalizar fenómenos binarios (sólo dos posibles concreciones), en la difusión del modelo en aplicaciones econométricas, como en la introducción de transformaciones del modelo, como por ejemplo la creación del Modelo Logit Mixto, como un modelo alternativo que puede situarse entre el Logit (mantiene su simplicidad) y el Probit (comparte su flexibilidad).

El empuje definitivo para el reconocimiento de Logit Model lo encontramos en los trabajos del economista americano McFadden, quien vincula este modelo a la Teoría de la Elección Discreta, abriendo un nuevo campo de trabajo que le hizo merecedor del Premio Nóbel de Economía en el año 2000. Los trabajos de McFadden se remontan a 1973, cuando trabajaba en California como consultor en un proyecto público en materia de transporte. Este economista fue pionero en usar el modelo logit para representar las preferencias de los individuos.

La importancia de los modelos de elección discreta radica en que permiten la modelización de variables cualitativas, característica que exige la codificación de la variable como paso previo a la modelización. En este proceso, los distintos estados de la variable se transforman en códigos o valores susceptibles de ser tratados utilizando técnicas de regresión.

McFadden planteó inicialmente el caso en el que los individuos se enfrentan a procesos de decisión dicotómicos, es decir, en los que únicamente hay dos posibles alternativas que representa de la forma y = 1 o y = 0.

El argumento principal de su tesis era que cada individuo tiene una función de utilidad U_i asociada a cada una de las alternativas ($y=0;\ y=1$). Esta función de utilidad puede dividirse en una componente sistemática V_i, que recoge el efecto de las variables explicativas (atributos observables), y una componente aleatoria ε_i, que recoge los efectos que tanto las variables no relevantes individualmente como del azar pueden tener sobre la utilidad. Según esto, la función de utilidad de un individuo se representa por la expresión

$$U_i = V_i + \varepsilon_i \,. \tag{9}$$

Si aceptamos la hipótesis de linealidad para la componente sistemática, la función de utilidad para cada estado posible de la elección se formula en los siguientes términos:

$$\begin{array}{l} \text{para } y = 0 \rightarrow U_0 = \alpha_0 + \beta_0 X + \varepsilon_0 \\ \text{para } y = 1 \rightarrow U_1 = \alpha_1 + \beta_1 X + \varepsilon_1 \end{array}, \tag{10}$$

donde α y β son parámetros y ε_0 y ε_1 se supone son independientes e idénticamente distribuidos (i.i.d.), según el modelo de valores extremo tipo I de Gumbel.

Si el comportamiento del individuo obedece al principio económico de maximización, elegirá la alternativa que le proporcione la máxima utilidad. En esta situación, la probabilidad de que el individuo elija la alternativa representada como y=1 será:

$$\begin{array}{l} P(y=1) = P(U_0 < U_1) = P\big(\varepsilon_0 - \varepsilon_1 < (\alpha_1 - \alpha_0) + (\beta_1 - \beta_0)X\big) = \\ \qquad = F\big[(\alpha_1 - \alpha_0) + (\beta_1 - \beta_0)X\big] \end{array}, \tag{11}$$

siendo F la función de distribución.

Llegado este punto, el problema se centra en el modelo de probabilidad de F. McFadden demostró que en las condiciones descritas $(\varepsilon_1 \text{ y } \varepsilon_0 \text{ i.i.d.})$ la variable diferencia de términos error $(\varepsilon_0 - \varepsilon_1)$ sigue un modelo de distribución de probabilidad logística.

Las innovaciones de McFadden proporcionaron al Modelo Logit una sólida base teórica, al tiempo que le dotó de un significado propio como modelo de elección discreta.

Por último, señalar que los últimos avances en el campo de la regresión tienden a unificar la teoría existente en cuanto a Modelos Probit y Logit con los modelos lineales generalizados basados en la distribución Normal y el análisis de la varianza. La tendencia en el siglo XXI no es la competición o exclusión entre modelos de regresión, sino una búsqueda del modelo más adecuado a las características del fenómeno observado, sean cuantitativas o cualitativas, proponiendo, incluso, una mixtura entre ellos.

Bibliografía

AITCHISON, J.; BROWN, J. A. (1957). *The lognormal Distribution*. Number 5 in University of Cambridge, Department of Applied Economics Monographs. Cambridge University Press, Cambridge.

BERKSON, J. (1944). “Application to the Logistic Function to Bio-assay”. Journal of The American Statistical Association, 39, 357-365.

BLISS, C. I. (1934). “The method of probit”. Science, 79, 38-39.

CRAMER, J. S. (2002). *Logit Models. From Economics and Other Fields*. Cambridge University Press, Cambridge.

CRAMER, J. S. (2003). *The Origins of Logistic regresion*. Tinbergn Institute Working Paper, nº 2002, 119/4

CORNFIELD, J. (1951). "A method of estimating comparative rates from clinical data". Journal of the National Cancer Institute, 11, 1269-1275.

COX, D. R. (1958). "The regresion analysis of binary sequences". Journal of the Royal Statistical Society, Series B, 20, 215-242.

COX, D. R. (1958). *The Analysis of Binaria Data*. Chapman and Hall. London

DU PASQUIER, L.-G. (1918). "Esquisse d'une nouvelle théorie de la population". Vierteljahrssshrift der Naturforschenden Gesellschaft, 63, 236-249.

KINGSLAND, S. E. (1985). *Modeling Nature*. The University of Chicago Press, Chicago.

LANDAV, D.; M LAZARSFELD, P. F. (1976). "Quetelet, Adolphe". Enciclopedia Internacional de las Ciencias Sociales. Aguilar. Madrid.

LIAGRE, J. B. (1879). *Calcul des probabilités et théorie des erreurs*. Bruselas: Muquardt. 2ª edición.

LOTKA, A. J. (1925). *Elementos de la Biología Física*. Williams y Wilkins publicaciones, Baltimore.

MALTHUS, T. E. (1798). *An Essay on the Principle of Population*. London.

MCFADDEN, D. (1974). "Condicional Logit análisis of Qualitative Choice Behavior". En Zerenbka (ed.). *Fronties in Econometrics*. New York.

MCFADDEN, D. (2001). "Economic choice". American Economic Review, 91, 352-357. (Discurso de aceptación del Premio Nobel)

PEARL, R. L.; REED, L. J. (1920). "On the rate of growth of the population of the United status since 1870 and its mathematical representation". Proceedings of the National Academy of Sciences, 6, 275-288.

PEARL R. (1925). *La biología del crecimiento de la población*. Knopf, New York.

REED, L. J.; BERKSON, J. (1929). "The application of the logistic function to experimental data". Journal of Physical Chemistry, 33, 760-779.

STIGLER, S. M. (1986). *The History of Statistics*. Cambridge, Mass. Harvard University Press.

VERHULST, P-F. (1838). *Notice sur la loi que la population suit dans son accroissement. Correspondance Mathématique et Physique*. Publicado por A. Quetelet, 10, 113.

VERHULST, P-F. (1845). "La Loi d'Accoissemrnt de la Population". Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-lettres de Belgique, 18, 1-59.

VERHULST, P-F. (1847). "La Loi d'Accoissemrnt de la Population". Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-lettres de Belgique, 20, 1-32.

YULE, G. U. (1925). "The growth of population and the factors which control". Journal of Royal Statistical Society, 138, 1-59.

Capítulo 31

Breve historia de la familia de clasificadores boosting

ESTEBAN ALFARO CORTES
MATIAS GAMEZ MARTINEZ
NOELIA GARCIA RUBIO
JOSE LUIS ALFARO NAVARRO
Universidad de Castilla-La Mancha. Albacete
JOSÉ MONDÉJAR JIMÉNEZ
Universidad de Castilla-La Mancha. Cuenca

Introducción

El aprendizaje automático incluye técnicas de clasificación para aprender a hacer predicciones precisas en base a lo que se ha observado en el pasado. Por ejemplo, para la predicción del fracaso empresarial intentamos construir un clasificador capaz de distinguir las empresas sanas de aquellas que van a fracasar. El planteamiento del aprendizaje automático ante este problema sería el siguiente: Empieza recopilando tantos ejemplos como sea posible de empresas que fracasaron así como de empresas sanas. A continuación con estas observaciones y las etiquetas con sus clases se entrena al sistema de aprendizaje elegido que produce un clasificador o regla de predicción. Posteriormente, ante una nueva empresa de clase desconocida, dicho clasificador intenta predecir si fracasará o no. El objetivo, por supuesto, es generar un clasificador que realice las predicciones más precisas posibles en nuevos ejemplos de test.

Construir un clasificador con una precisión muy alta es en ocasiones una tarea muy difícil. Sin embargo, no es difícil conseguir reglas sencillas con una precisión mucho más moderada. Un ejemplo de este tipo de reglas podría ser algo como lo siguiente: “Si la empresa presenta pérdidas continuadas en los tres últimos años, entonces asignarla al grupo de empresas que va a fracasar”. Una regla así ni siquiera cubre todas las empresas, ya que no dice nada de aquellas empresas que no tengan pérdidas continuadas en los tres últimos ejercicios. Sin embargo, las predicciones de esta regla son mejores que las tomadas sólo en base al azar.

Boosting, el método de clasificación que constituye el objeto de este trabajo, se basa en el hecho de que encontrar muchas reglas sencillas puede ser mucho más fácil que encontrar una única regla con una alta precisión para la predicción. Para aplicar el método boosting, empezaremos eligiendo el algoritmo para encontrar las reglas sencillas. El algoritmo boosting entrena a estos clasificadores básicos o débiles repetidas veces, utilizando cada vez distintos conjuntos de entrenamiento, o de forma más precisa, distintas distribuciones o pesos sobre el conjunto de entrenamiento. Cada vez que se le llama, el clasificador básico genera una nueva regla de predicción débil, y después de muchas iteraciones, el algoritmo boosting debe combinar estas reglas débiles en una única regla de predicción que, se espera, sean mucho más precisa que ninguna de las reglas débiles.

Para utilizar este procedimiento hay que tomar dos decisiones fundamentales: primero, cómo debe seleccionarse la distribución en cada iteración, y segundo, cómo deben combinarse las reglas básicas en una sola. Respecto a la elección de la distribución, la técnica utilizada es aumentar el peso de las observaciones que han sido erróneamente clasificadas más a menudo por los clasificadores básicos de las iteraciones anteriores, de esta forma forzamos al clasificador base a centrar su atención en los ejemplos más difíciles. Mientras que para combinar los clasificadores básicos, simplemente tomar el voto mayoritario (ponderado) de sus predicciones es una solución natural y efectiva. Una tercera elección también importante es el método de clasificación que se utiliza para generar los clasificadores sencillos, pero en este trabajo dejaremos abierta esta cuestión ya que se pueden utilizar árboles de clasificación, redes neuronales o cualquier otro, sin que esto sea importante para la descripción general de este método.

Por tanto, Boosting es un método general y probablemente efectivo de producir una regla de predicción precisa combinando reglas sencillas y moderadamente imprecisas de forma similar a la descrita anteriormente. En este trabajo presentamos una revisión histórica de algunos de los trabajos recientes en boosting, centrándonos especialmente en el algoritmo AdaBoost que ha sufrido un intenso estudio teórico y comprobaciones empíricas.

Los orígenes del Boosting

En el aprendizaje supervisado, a partir de un conjunto de entrenamiento dado, el sistema de clasificación generará una regla que asigna a cada observación una de las posibles clases. Esta regla será más o menos precisa dependiendo de la calidad del sistema de clasificación y de la dificultad del problema en cuestión. Intuitivamente, se puede ver que si el clasificador se comporta mejor que la regla por defecto, aunque sólo sea ligeramente, querrá decir que el sistema de clasificación ha encontrado cierta estructura en los datos para lograr su ventaja. Boosting es un método que potencia[1] la precisión del sistema de clasificación aprovechando su ventaja.

Boosting utiliza el sistema de clasificación como una subrutina para generar una regla de predicción, que tiene asegurada una elevada precisión en el conjunto de entrenamiento. Este método aplica el sistema de clasificación en sucesivas ocasiones sobre el conjunto de entrenamiento, pero cada vez centra la atención en ejemplos distintos.

Una vez finalizado el proceso, se combinan los clasificadores básicos en un único clasificador final muy preciso en el conjunto de entrenamiento. Además, este clasificador final, normalmente, es también muy preciso en el conjunto de prueba. Es decir, obtiene

[1] En realidad la traducción literal de Boosting es potenciando.

también notables resultados en la generalización, como han comprobado diversos autores tanto teórica como empíricamente.

Aunque los orígenes pueden remontarse un poco más allá, donde se unen el campo del reconocimiento de patrones y la inteligencia artificial, generalmente se considera que el primer algoritmo de lo que hoy en día ya es una familia de algoritmos Boosting, fue el que propuso Robert E. Schapire en 1990 en su artículo "*The strength of weak learneability*". En ese artículo se propone un algoritmo para potenciar la precisión de lo que se denomina clasificador débil, aquél que es sólo ligeramente superior a la regla del azar o regla por defecto. Este algoritmo es adecuado, en principio, para el caso dicotómico, en el que sólo existen dos clases y se supone que el clasificador C_1 construido sobre el conjunto de entrenamiento T de tamaño n, es sólo ligeramente superior a la regla por defecto, es decir, su error es $\varepsilon = 0,5 - \gamma$, donde γ es la ventaja que el clasificador C_1 obtiene sobre la regla por defecto.

Para forzar al sistema de clasificación a aprender más sobre las observaciones difíciles de la distribución, se debe eliminar la ventaja del clasificador C_1 de alguna manera. Schapire propone, a partir del conjunto de entrenamiento T o T_1, construir un nuevo conjunto de entrenamiento T_2, del mismo tamaño n, forzando a que en este nuevo conjunto el clasificador C_1 no sea mejor que la regla por defecto, es decir, que la mitad de los ejemplos en T_2 estén bien clasificados por C_1 pero se equivoque en la otra mitad. Para lograr esto, se lanza una moneda al aire, si sale cara se extraen aleatoriamente ejemplos de T hasta encontrar uno que C_1 sea capaz de clasificar correctamente, es decir, $C_1(x_i) = y_i$. Y si sale cruz, se extraen ejemplos de T hasta que salga uno mal clasificado, o sea, $C_1(x_i) \neq y_i$.

El autor explica que el tiempo de búsqueda es limitado, ya que si la precisión del clasificador C_1 está muy próxima a 0'5, el número de ejemplos extraídos hasta encontrar uno que cumpla con el requisito deseado, estar bien o mal clasificado según salga cara o cruz, sigue una distribución geométrica con probabilidad aproximadamente igual a 0,5. La esperanza de una distribución geométrica de parámetro p= 0'5, es $(1-p)/p = 1$. Por tanto, en promedio, se puede esperar que sólo sea necesario extraer dos observaciones, la deseada y otra.

Una vez formado el conjunto de entrenamiento T_2, se entrena de nuevo el sistema de clasificación pero utilizando, en esta ocasión, para ello el conjunto T_2, se obtiene así el clasificador C_2 ,cuyo error debe ser también al menos ligeramente inferior al de la regla por defecto. Por último, se crea un tercer conjunto T_3 , de tamaño n, eliminando de T aquellos ejemplos donde C_1 y C_2 coinciden. Es decir, se extraen observaciones de T hasta encontrar una en la que $C_1(x_i) \neq C_2(x_i)$. Se entrena por tercera vez al sistema de clasificación creando el clasificador C_3.

Para clasificar una nueva observación, x_i, si $C_1(x_i)=C_2(x_i)$ entonces se le asigna la clase acordada por ambos clasificadores, en los demás casos, C_F asigna $C_3(x_i)$. Es decir, en la combinación final C_F se aplica el voto mayoritario de C_1 , C_2 y C_3. Schapire demostró que el error del C_F está acotado por $g(\varepsilon)= 3\varepsilon^2-2\varepsilon^3$, que es significativamente menor que el error original ε. Este proceso supone el núcleo fundamental del algoritmo boosting, y se utiliza reiteradamente para mejorar la precisión del clasificador final, llegando a formarse una estructura un tanto complicada que podría reflejarse en la Figura 1.

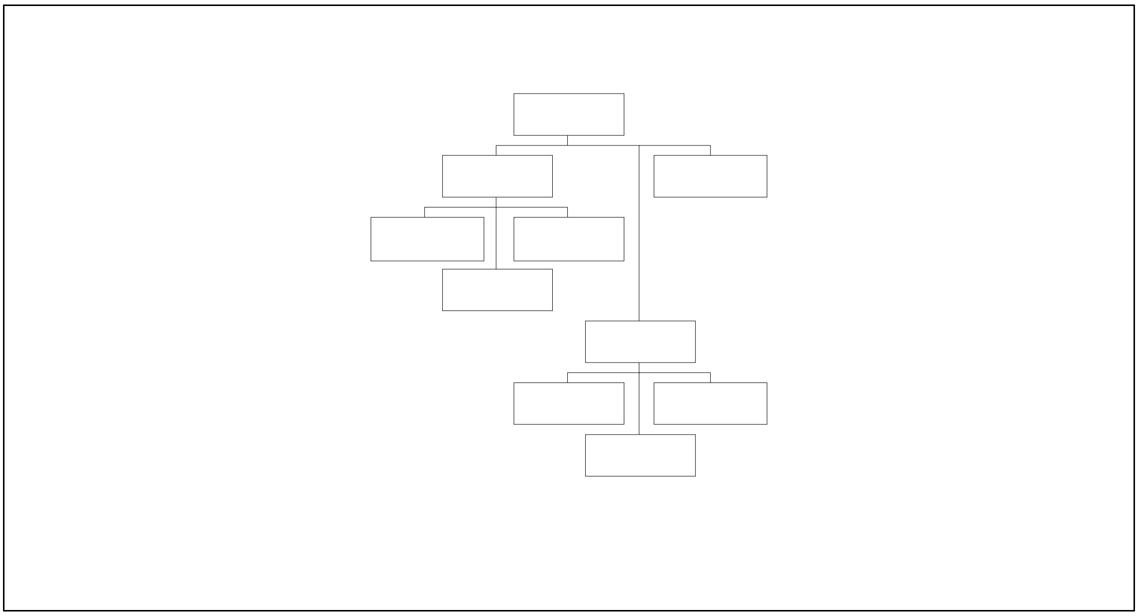


Figura 1: Estructura del algoritmo de Schapire (1990)

A pesar de la importancia que tuvo este primer algoritmo de Schapire su aplicación en la práctica resulta algo complicada. Además, la forma de construir los nuevos conjuntos es poco eficiente, debido a las llamadas recursivas que debe realizar. Poco tiempo después del trabajo de Schapire, Yoav Freund desarrolló un algoritmo más sencillo y eficiente que presentó en su artículo *"Boosting a weak learning algorithm by majority"* Este algoritmo también construye varias distribuciones diferentes sobre el conjunto de observaciones, que se presentan al sistema de clasificación para centrar su atención en las regiones difíciles de la distribución desconocida. El sistema de clasificación genera un clasificador básico para cada distribución que recibe. Por tanto, los clasificadores básicos se comportarán bien en diferentes regiones del espacio de observaciones. El algoritmo boosting combina los clasificadores básicos en un clasificador final mediante el voto mayoritario simple.

En cada iteración, b=1,2, .., B, a los ejemplos del conjunto de entrenamiento se les asigna una distribución de probabilidad en función de la cual se genera el clasificador básico correspondiente a la iteración b-ésima. Aquellos ejemplos clasificados correctamente por ese clasificador reciben una marca. Después de las B iteraciones, los ejemplos que han sido marcados más de B/2 veces, son los ejemplos clasificados correctamente por el clasificador final, C_F. Las observaciones restantes, aquellas que han sido marcadas B/2 veces o menos[1], serán clasificadas incorrectamente por el clasificador final, el conjunto de estas observaciones se representa por L. Por tanto, el error de C_F en el conjunto de entrenamiento es |L| / |T|. El objetivo de este algoritmo boosting es minimizar este error. Para minimizar este error, Freund propone una estrategia de ponderación, que en cada iteración actualiza el peso de la observación x_i en la iteración b-ésima en función de b, B, γ y de cuántas veces x_i ha sido marcada antes de esa iteración. Donde, como ya se ha dicho, γ es la ventaja que los clasificadores básicos deben obtener como mínimo sobre la regla por defecto, siendo su error $\varepsilon_b = 0{,}5-\gamma$. Para más detalles sobre la estrategia concreta de actualización de los pesos en este algoritmo, véase Freund (1995).

[1] En este caso, se trabaja bajo el supuesto pesimista de que si los empates se rompen de forma aleatoria, la clase elegida es siempre errónea. Por ello, cuando una observación ha sido marcada exactamente B/2 veces, se considera que el clasificador final se equivoca.

Lo más importante del algoritmo de Freund es que consigue una estructura más sencilla que la que planteó Schapire originalmente, manteniendo e incluso mejorando la eficiencia, ya que consigue reducir el error de entrenamiento más rápidamente. La Figura 2 ayuda a comprender mejor la diferencia de la estructura de ambos algoritmos.

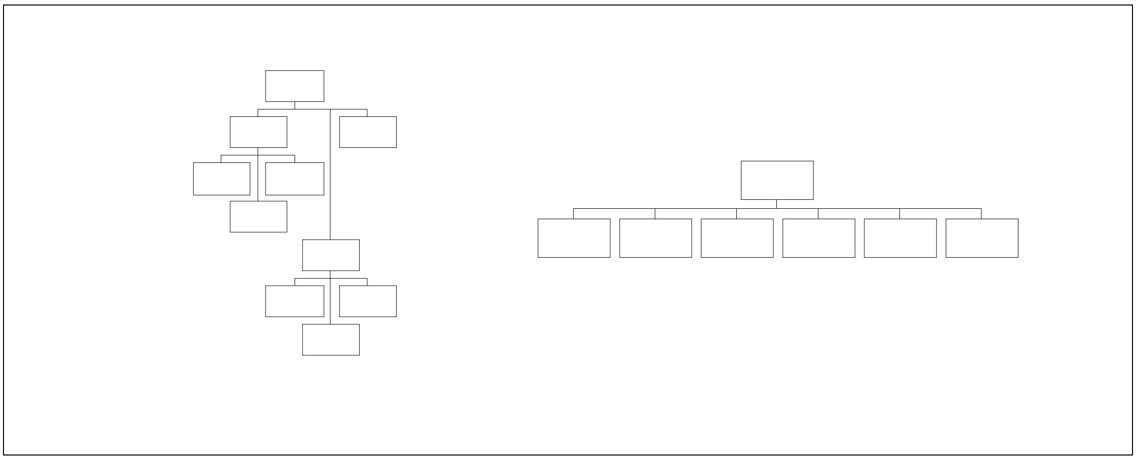


Figura 2. La parte izquierda del gráfico muestra la estructura compleja del algoritmo de Schapire (1990). A la derecha, la estructura más sencilla de Freund (1995).

Sin embargo, este algoritmo presenta también algunas deficiencias prácticas. En primer lugar, la regla de actualización de los pesos depende de γ, la ventaja que como mínimo deben obtener los clasificadores básicos sobre la regla por defecto. Es decir, esta es la ventaja que obtendrá en el peor de los casos, pero es posible que en realidad se obtenga una ventaja superior en muchos de los casos. En segundo lugar, Freund probó que el algoritmo boosting por mayoría requiere aproximadamente $1/\gamma^2$ iteraciones para reducir el error de entrenamiento hasta cero. Lo que quiere decir que, en el caso extremo de que γ=0,001, se necesitarán un millón de iteraciones. Durante las B iteraciones del proceso, algunos de los clasificadores básicos obtendrán un error menor que 0,5 - γ, pero este algoritmo no es capaz de aprovechar esta ventaja para acelerar el proceso.

Poco tiempo después, Freund y Schapire desarrollaron el algoritmo boosting más utilizado, el AdaBoost (Adaptative boosting), que presentaron en el artículo *"Experiments with a new Boosting Algorithm"* publicado en 1996. En este caso, en la iteración b-ésima se aumenta el peso de los ejemplos mal clasificados por el clasificador básico correspondiente a esa iteración. Además, la regla de actualización de los pesos ya no depende de la constante γ, sino del error obtenido por el clasificador básico en esa iteración.

El comportamiento general de AdaBoost es similar al de los otros algoritmos boosting, puesto que los clasificadores básicos se generan de forma sucesiva y se centra la atención en los ejemplos que resultan más difíciles. La principal diferencia entre AdaBoost y el algoritmo de boosting por mayoría es la regla de actualización de los pesos. AdaBoost utiliza una regla de actualización que depende del error del clasificador básico actual, no de la ventaja que deben conseguir como mínimo, γ. Es decir, en AdaBoost ese factor de la regla de actualización varía en cada iteración, mientras que en el otro caso permanece constante.

Además, otra diferencia es que a cada clasificador básico se le atribuye un peso α_b en función del error que comete. Este peso se utiliza en la combinación final, puesto que AdaBoost utiliza el voto mayoritario ponderado de los clasificadores básicos, en lugar del voto mayoritario simple. De esta forma, aquellos clasificadores básicos que consiguen una mayor precisión tendrán una mayor importancia en la combinación final.

El algoritmo AdaBoost ha facilitado la aplicación del método boosting y desde su aparición han sido mucho los trabajos realizados al respecto, tanto empíricos comprobando sus buenos resultados, como teóricos, intentando explicar por qué se obtienen tan buenos resultados. Como resultado de esta extensa labor investigadora han surgido diversas modificaciones a este algoritmo, algunas de las cuales se recogen más adelante en el apartado 4 de este trabajo.

AdaBoost

Aunque existen diversas versiones de algoritmos boosting la más extendida es la que proporcionan Freund y Schapire (1996) que se conoce como AdaBoost. Para simplificar se puede suponer que sólo existen dos clases y en el capítulo siguiente se generaliza a más de dos clases. Se parte del conjunto de entrenamiento T = $\{(x_1^b, y_1^b), ..., (x_n^b, y_n^b)\}$ donde Y toma en este caso los valores$\{-1, 1\}$. Se asigna a cada observación x_i el peso $\omega_b(i)$, que inicialmente se iguala a 1/n, y posteriormente se irá actualizando en cada iteración. Se construye un clasificador básico a partir del conjunto de entrenamiento ponderado, que se representa por $C_b(x_i)$ y se aplica a cada uno de los ejemplos de entrenamiento. El error de ese clasificador se representa por ε_b y se calcula como

$$\varepsilon_b = \sum_{i=1}^{n} \omega_b(i)\xi_b(i) \qquad \text{donde} \qquad \xi_b(i) = \begin{cases} 0 & C_b(x_i) = y_i \\ 1 & C_b(x_i) \neq y_i \end{cases}, \tag{1}$$

A partir del error del clasificador en la iteración *b*-ésima, se calcula la constante α_b, que se utiliza para la actualización de los pesos. En concreto estos autores hacen $\alpha_b = \ln(1 - \varepsilon_b / \varepsilon_b)$ y el nuevo peso para la iteración *b+1* será

$$\omega_{b+1}(i) = \omega_b(i) \cdot \exp(\alpha_b \xi_b(i)). \tag{2}$$

Posteriormente se normalizan los pesos calculados para que la suma de todos ellos sea uno. Según Freund y Schapire, el error debe ser inferior al de la regla por defecto, $\varepsilon_b = 0,5 - \gamma_b$, donde γ_b representa la ventaja que obtiene el clasificador básico de la iteración *b*-ésima sobre la regla por defecto, en el peor de los casos en que las dos clases tengan la misma probabilidad a priori, 0'5.

Como se puede ver, aumenta la ponderación de las observaciones mal clasificadas y disminuye la de las clasificadas correctamente, forzando así al clasificador básico construido en la siguiente iteración a centrarse en aquellos casos que han resultado más difíciles. Además, las diferencias en la actualización son mayores cuando el error cometido por el clasificador básico es pequeño, porque si el clasificador consigue una precisión elevada se le da más importancia a los pocos fallos cometidos. Por tanto, la constante alfa puede considerarse como una tasa de aprendizaje calculada en función del error cometido en esa iteración. Además, esta constante también se utiliza en la regla de decisión final, dando más importancia a los clasificadores básicos que cometen un menor error.

Este proceso se repite en todas las iteraciones desde b =1, 2, 3, ..,B. Para acabar, se construye el clasificador final como combinación lineal de los clasificadores básicos ponderados por la constante α_b correspondiente. En concreto será:

$$C(x) = sign\left(\sum_{b=1}^{B} \alpha_b C_b(x)\right) \qquad \mathbf{(3)}$$

A continuación mostramos El algoritmo AdaBoost

Algoritmo 1. AdaBoost (Freund y Schapire, 1996)
1. Iniciar con $\omega_1(i) = 1/n$, i=1, 2, ..., n. 2. Repetir para b=1, 2, ...,B a) Construir el clasificador $C_b(x) \in \{-1,1\}$ utilizando los pesos $\omega_b(i)$ en T^b. b) Calcular: $\varepsilon_b = \sum_{i=1}^{n} \omega_b(i)\xi_b(i)$ y $\alpha_b = ln(1 - \varepsilon_b/\varepsilon_b)$ c) Actualizar los pesos $\omega_{b+1}(i) = \omega_b(i) \cdot \exp(\alpha_b \xi_b(i))$ y normalizarlos. 3. Construir el clasificador final $C(x) = sign\left(\sum_{b=1}^{B} \alpha_b C_b(x)\right)$

AdaBoost puede aplicarse de dos maneras distintas, utilizando remuestreo o utilizando reponderación. En la versión que utiliza remuestreo, se obtiene el conjunto de datos S_b para la iteración b-ésima, mediante una submuestra bootstrap extraída con reemplazamiento, utilizando como probabilidades para la extracción los pesos de las distintas observaciones en esa iteración. En la versión que utiliza la reponderación, el clasificador C_b tiene en cuenta directamente los pesos de los ejemplos. No existe una evidencia fuerte a favor de ninguno de los métodos frente al otro (Breiman,1998), (Freund y Schapire, 1997) y (Freund y Schapire, 1998)

Versiones de AdaBoost

El algoritmo AdaBoost descrito en el apartado 6.4, entrena varias veces al clasificador en el mismo conjunto de entrenamiento, pero centrando su atención en cada paso en los ejemplos que han sido mal clasificados en el paso anterior. AdaBoost se limita al caso dicotómico y utiliza clasificadores básicos cuya salida puede tomar únicamente los valores -1 y 1, por ello se conoce también como AdaBoost discreto, a partir del trabajo de Friedman y otros (2000). Desde su aparición en 1996, han sido varias las modificaciones propuestas para mejorar los algoritmos boosting, en este trabajo se recogen algunas de ellas.

Una generalización del AdaBoost discreto fue propuesta en Freund y Schapire (1996) y explicada en mayor profundidad en Schapire y Singer (1999), esta generalización utiliza predicciones con valores reales, que expresan además el grado de confianza, en lugar de los valores {-1, 1} que utiliza AdaBoost. En este caso, los clasificadores básicos asignan un valor real entre -1 y 1. De tal forma que el signo de $C_b(x)$ informa de la clase predicha y el valor absoluto, $|C_b(x)|$ da una medida de la confianza en la predicción. El valor real de esta contribución se combina con las anteriores utilizando, como en el algoritmo original, un coeficiente α_b, aunque en esta ocasión se calcula de forma ligeramente distinta.

Friedman(2000) presenta una versión que llama AdaBoost Real, donde los clasificadores básicos estiman la probabilidad de pertenencia a una determinada clase $p_b(x)=P_\omega(Y=1/X=x)\in[0,1]$, donde el subíndice ω indica que esa probabilidad se calcula sobre el conjunto ponderado utilizando los pesos $\omega_b(i)$ correspondientes a cada iteración. La contribución al clasificador final es la mitad de la transformación logit de esta probabilidad estimada.

Algoritmo 2. AdaBoost Real.
1. Iniciar con $\omega_1(i) = 1/n$, i=1, 2, ..., n. 2. Repetir para b=1, 2, ...,B a) Construir un clasificador para obtener una estimación de las probabilidades de las clases $p_m(x) = \hat{P}_\omega(y=1/x) \in [0,1]$ utilizando los pesos $\omega_b(i)$ en T^b. b) Calcular: $f_b(x) = 1/2 \ln(p_m(x)/(1-p_m(x))) \in R$ c) Actualizar los pesos $\omega_{b+1}(i) = \omega_b(i)\cdot\exp(-y_b f_b(x_i))$ y normalizarlos. 3. Construir el clasificador final $C(x) = sign\left(\sum_{b=1}^{B} f_b(x)\right)$

Como se ha visto en el apartado anterior, el algoritmo AdaBoost minimiza una función de pérdida exponencial. Sin embargo, diversos autores defienden que para clasificación, en lugar de una función de pérdida exponencial, resulta más natural la elección como función de pérdida del logaritmo de la verosimilitud binomial. Basándose en esta función Friedman (2000) propone un nuevo algoritmo que llama Logitboost. Este nuevo algoritmo mantiene la misma filosofía de los anteriores algoritmos boosting, entrenar en sucesivas ocasiones al clasificador forzándolo a centrarse en los ejemplos difíciles, pero la forma de llevarlo a cabo es algo diferente.

En cada iteración se crea una función $f_b(x_i)$ que se ajuste lo mejor posible a z_i^b, que es la respuesta objetivo con la que se trabaja en la iteración b-ésima para la observación i-ésima. Esta respuesta se calcula en función de la verdadera clase de esa observación, que en este caso será {0,1}, de la probabilidad que se le ha asignado a la clase 1 en la iteración anterior, y del peso en esa iteración.

$$z_i^b = \frac{y_i - p_{b-1}(x_i)}{\omega_i^b} \qquad \textbf{(4)}$$

donde $p_{b-1}(x_i) = P_\omega(Y=1/X=x_i)$ y $\omega_b(i)= p_{b-1}(x_i)\cdot(1-p_{b-1}(x_i))$.

$F_b(x_i)$ es una función de agregación que se va actualizando en cada iteración y se calcula como $F_b(x_i) = F_{b-1}(x_i) + 0{,}5\, f_b(x_i)$ y a partir de ella se calcula la probabilidad de que la observación i-ésima pertenezca a la clase 1 en la iteración b.

$$p_b(x_i) = \frac{\exp(F_b(x_i))}{\exp(F_b(x_i)) + \exp(-F_b(x_i))} = \frac{1}{1+\exp(-2F_b(x_i))} \quad \textbf{(5)}$$

Para comprender mejor este algoritmo cada función de agregación, $F_b(x_i)$, puede considerarse como una estimación de la mitad del logaritmo neperiano del cociente de probabilidades de las clases.

$$F(x) = \frac{1}{2}\ln\left(\frac{p(x)}{1-p(x)}\right) \quad \textbf{(6)}$$

Por tanto, Logitboost ajusta un modelo de regresión logística aditivo mediante optimización secuencial del logaritmo de la verosimilitud binomial, para más detalles véase Friedman (2000).

Una propiedad muy útil de este método es que produce directamente estimaciones de las probabilidades $\hat{P}$[Y=1/ X=x]. Esto es muy importante para construir clasificadores cuando los costes de error en la clasificación no son iguales. Además, permite construir clasificadores con la opción de no asignar ninguna clase o asignar la etiqueta "duda" o "no clase" para ciertas observaciones. Una ventaja importante de Logitboost comparado con métodos como las redes neuronales es que funciona bien sin necesidad de un ajuste afinado y sin llevar a cabo una sofisticada optimización no lineal, aunque esta ventaja es general para los algoritmos boosting.

Algoritmo 3. LogitBoost.
1. Iniciar con una función de agregación $F_0(x) \equiv 0$ y probabilidades $p_0(x) \equiv ½$, donde $p(x)$ es la forma abreviada de $\hat{P}$[Y=1/X=x] 2. Repetir para b=1, 2, ...,B a) Calcular la respuesta de trabajo y los pesos para i=1,2, ..,n $\omega_b(i) = p_{b-1}(x_i)(1-p_{b-1}(x_i))$; $z_i^b = \frac{y_i - p_{b-1}(x_i)}{\omega_i^b}$ b) Ajustar una función $f_b(x_i)$ por mínimos cuadrados ponderados $f_b(x_i) = \arg\min_f \sum_{i=1}^{n} \omega_b(i)(z_i^b - f_b(x_i))^2$ c) Actualizar la función de agregación y calcular las nuevas probabilidades. $F_b(x_i) = F_{b-1}(x_i) + 0{,}5\, f_b(x_i)$ $p_b(x_i) = (1+\exp(-2\,F_b(x_i)))^{-1}$ 3. Construir el clasificador final. $C(x_i) = sign(F_b(x_i))$

Dada la definición de los pesos $\omega(x)$ en aquellos ejemplos donde $p(x)$ esté próximo a 0 ó a 1, $\omega(x)$ llegará a ser muy pequeño. Esto puede causar problemas en el cálculo de $z(x_i)$, por lo que en Friedman(2000) se aconseja tomar las siguientes precauciones:

Si y=1, entonces calcular $z= (y-p)/p(1-p)$ como $1/p$. Como este cociente puede hacerse muy grande si p es pequeño, se debe limitar su valor en una cantidad determinada que se puede representar por *zmax*. El valor concreto elegido para *zmax* no es crucial y Friedman(2000) afirma que valores comprendidos entre 2 y 4 funcionan bien. Por otro lado, si y=0, calcular z como $-1/(1-p)$ con el límite inferior de *-zmax*.

Leo Breiman, autor del método Bagging (Breiman, 1996), mostró inmediatamente un gran interés por el método Boosting que Freund y Schapire propusieron casi simultáneamente a Bagging y que obtenía mejores resultados que éste. El considera que lo fundamental de esta técnica es el uso de remuestreo adaptativo y combinación, y por eso, los llama algoritmos Arcing[1]. La principal diferencia con el método Bagging consiste en entrenar los clasificadores básicos en muestras generadas a partir de distribuciones de probabilidad que van cambiando en función de los errores cometidos, en lugar de permanecer constantes como ocurre en Bagging. Esto implica que los clasificadores básicos utilizados en Bagging son independientes, mientras que en el caso de Boosting cada clasificador depende de los anteriores.

En Breiman (1998) el autor explica que después de probar AdaBoost empezó a sospechar que el éxito de éste no radicaba en su forma concreta sino en el remuestreo adaptativo que realiza, donde se aumentan los pesos de aquellos ejemplos que son clasificados erróneamente con mayor frecuencia. Para comprobarlo, probó tres maneras sencillas de actualizar los pesos o probabilidades. En todos ellos la actualización se realiza en función del valor $1 + m_i^{\,h}$, siendo m_i el número de veces que la observación *i*-ésima ha sido clasificada incorrectamente por los clasificadores básicos construidos hasta esa iteración. Breiman utilizó los valores h=1,2, 4 y este último fue el que mejores resultados le proporcionó. Otra diferencia de arcing respecto a Boosting es que no utiliza ponderaciones en la regla de combinación final, lo que supone también una mayor sencillez del algoritmo al no tener que calcular y guardar esas ponderaciones.

Algoritmo 4. Arcing
1. Iniciar con $\omega_1(i) = 1/n$, i=1, 2, ..., n. 2. Repetir para b=1, 2, ...,B a) Construir el clasificador $C_b(x)$ utilizando los pesos $\omega_b(i)$ para extraer con reemplazamiento la muestra aleatoria T^b. b) Clasificar los ejemplos de T utilizando $C_b(x)$ y calcular $m_b(x_i)$ que es el número de errores cometidos por los b primeros clasificadores en la observación i-ésima. c) Actualizar los pesos $w_{b+1}(i) = \dfrac{1+m_b(x_i)^4}{\sum_{i=1}^{n} 1+m_b(x_i)^4}$ y normalizarlos. 3. Construir el clasificador final $C(x) = sign\left(\sum_{b=1}^{B} \alpha_b C_b(x)\right)$.

[1]Arcing es el acrónimo del término inglés Adaptive Resampling and Combining. Breiman llama arc-fs al algoritmo AdaBoost, en honor a Freund y Schapire, y a la modificación que él propone arc-x4, pero en este trabajo cuando se hable de arcing se referirá a arc-x4.

Conclusiones

En esta revisión, hemos visto que han surgido muchas interpretaciones o formas de ver de AdaBoost. En primer lugar y más importante, AdaBoost es un algoritmo boosting genuino: teniendo acceso a un verdadero algoritmo de aprendizaje débil que siempre se comporte ligeramente mejor que la regla del azar en todas las distribuciones sobre el conjunto de entrenamiento, podemos probar límites arbitrariamente buenos en el error de entrenamiento y de generalización de AdaBoost.

A partir de la visión original, AdaBoost ha sido interpretado como un procedimiento basado en gradiente descendente funcional o como una aproximación a la regresión logística, entre otras cosas. Todas estas conexiones e interpretaciones han permitido que comprendamos mejor el boosting y han contribuido a su extensión en direcciones incluso más prácticas, tales como regresión logística y otros problemas de minimización de funciones de pérdida, a problemas de más de tres clases, a incorporar regularización y permitir la integración de conocimiento a priori.

Bibliografía

BAUER, E. Y KOHAVI, R. (1999): "An empirical comparison of voting classification algorithm: Bagging, boosting and variants". Machine Learning, vol. 36, pp 105-142.

BREIMAN, L. (1996): "Bagging predictors". Machine Learning, Vol 24, 2, pp.123-140.

BREIMAN, L. (1998). "Arcing classifiers". The Annals of Statistics, Vol 26, 3, pp. 801-849.

DIETTERICH, T.G. (2000): "Ensemble methods in machine learning". En *Multiple Classifier Systems*, Cagliari, Italia.

DRUCKER, H. Y CORTES, C. (1996): "Boosting decision trees". En D.S. Touretzky, M.C. Mozer, and M.E. Hasselmo, editors, *Advances in Neural Information Processing Systems 8: Proc. of NIPS'95*, vol. 8, pp. 479-485. The MIT Press.

FREUND, Y. (1995): "Boosting a weak learning algorithm by majority". Information and Computation, 121(2):256-285.

FREUND, Y. Y SCHAPIRE, R.E. (1996): "Experiments with a New Boosting Algorithm". En *Proceedings of the Thirteenth International Conference on Machine Learning*, pp. 148-156, Morgan Kaufmann.

FREUND, Y. Y SCHAPIRE, R.E. (1997): "A decision-theoretic generalization of on-line learning and an application to boosting". Journal of Computer and System Sciences, 55[1], pp.119-139.

FREUND, Y. Y SCHAPIRE, R.E. (1998): "Discussion of the paper arcing classifiers" de Leo Breiman. The Annals of Statistics, 26[3], pp.824-832.

FREUND, Y. Y SCHAPIRE, R.E., BARTLETT, P. Y LEE, W.S. (1998): "Boosting the margin: A new explanation for the effectiveness of voting methods". The Annals of Statistics, 26(5), pp. 1651-1686.

FRIEDMAN, J.; HASTIE, T. Y TIBSHIRANI, R. (2000): "Additive logistic regression: a statistical view of boosting". The Annals of Statistics, 38(2), pp. 391-293.

KUNCHEVA, L.I. (2004): *Combining pattern classifiers. Methods and algorithms*. Wiley

OPITZ, D. Y MACLIN, R. (1999): "Popular ensemble methods: An empirical study". Journal of Artificial Intelligence Research, 11, pp.169-198.

QUINLAN, J.R. (1996): "Bagging, boosting, and C4.5". En *Proceedings of the Thirteenth National Conference on Artificial Intelligence and the Eighth Innovative Applications of Artificial Intelligence Conference*, pp. 725-730, Menlo Park. AAAI Press / MIT Press.

SCHAPIRE, R.E. (1990): "The strength of weak learnability". Machine Learning, 5(2):197-227.

SCHAPIRE, R.E. Y SINGER, Y. (1999): "Improved boosting algorithms using confidence-rated predictions". Machine Learning, 37(3):297-336.

VALENTINI, G. Y MASULLI, F. (2002): "Ensembles of learning machines". En Marinaro, M. y Tagliaferri, R. (ed) *Neural Nets WIRN Vietri*, Series Lecture Notes in Computer Sciences, Springer-Verlag, Heidelberg, Alemania.

Capítulo 32

Investigaciones factoriales sobre la inteligencia Técnicas en la década entre 1940 a 1950: Aplicaciones a la selección de personal y a la orientación profesional

MARCELO PASCUAL FAURA
Universidad San Pablo-CEU de Madrid

Introducción

El objetivo de este trabajo es exponer de forma resumida las investigaciones sobre la estructura factorial de la Inteligencia técnica entre los años1940 y 1950, así como su aplicación práctica a la selección de personal y a la orientación profesional de los escolares.

Estas investigaciones fundamentalmente se llevaron a cabo en Inglaterra, Francia y Estados Unidos. En Inglaterra por una parte la Psicología militar acapara la atención de los psicólogos durante los años de la Segunda Guerra Mundial, centrados en tareas de selección y clasificación de reclutas para las distintas especialidades de las fuerzas armadas; por otra parte, pero simultáneamente, se desarrolla un proyecto de reconstrucción nacional para la postguerra, en el cual se incluye un programa educativo con tres tipos de enseñanza secundaria, en el que los psicólogos deben enfrentar la orientación educativa profesional a la edad de once años, para cada tipo de enseñanza secundaria..

En Francia destacan las investigaciones realizadas por Germaine Benyer en una línea similar, tanto conceptualmente como metodológicamente, a la británica, pero teniendo en cuenta los resultados obtenidos por la Escuela Americana.

En Estados Unidos cabe destacar por una parte los trabajos realizados sobre la Estructura factorial de la inteligencia técnica en las Fuerzas Aéreas Americanas –A.A.F.- centrados en la selección y clasificación de personal para desempeñar tareas técnicas y mecánicas de tipo bélico y en especial con tripulaciones aéreas, donde el principal método de investigación fue el análisis de las correlaciones, en la validación de tests y baterías, la regresión simple y múltiple, y el análisis factorial exploratorio para definir la naturaleza de la contribución de los tests en el pronóstico de criterio. Por otra parte, en 1945, se publican los resultados del Proyecto "USES" sobre Orientación Escolar y Profesional.

Escuela Inglesa

Por una parte, la Psicología militar, acapara la atención de los psicólogos durante los años de la Segunda Guerra Mundial; centrándose fundamentalmente en temas de selección y clasificación de los reclutas para las distintas especialidades de las Fuerzas Armadas Británicas (Vernon, P.E. (1949[a]); Vernon, P.E. (1947B); Vernon, P.E. et Parry, J.B. (1949) resumen detalladamente los trabajos realizados en las Fuerzas Armadas Británicas en esta área de la Psicología militar).

Simultáneamente en los años de la Segunda Guerra Mundial se desarrolla en Inglaterra un Proyecto de Reconstrucción Nacional para la postguerra, en el cual se incluía un proyecto educativo, con tres tipos de enseñanza secundaria: una enseñanza de tipo verbal o literario; otra de formación manual o mecánica, y una tercera consistente en un aprendizaje práctico de tipo cultural. Para llevar a cabo este Proyecto, dos psicólogos deberían afrontar la orientación y selección de alumnos a la edad de once años, para cada uno de estos tres tipos de enseñanza.

Las investigaciones realizadas hasta esta época establecían claramente que la "Aptitud Verbal" (V), constituía una habilidad importante en todos los trabajos que, normalmente, desempeñaban los jóvenes al finalizar la enseñanza secundaria y que podía medirse a la edad de once años. También se tenía constancia que el "Juicio Espacial" (K), era importante para la mayoría de los trabajos en los que solían emplearse los jóvenes al finalizar la Escuela Técnica, aunque en este caso no se había verificado si el Juicio Espacial" (K) podía medirse a esa edad. Por otra parte, la "Habilidad práctica o de Ejecución" (F), se sabía que no podía medirse hasta edades superiores a los once años.

Para poder fundamentar científicamente los procesos de orientación y selección de los adolescentes hacia los diversos tipos de enseñanza secundaria proyectados, se inician una serie de investigaciones que tratan de averiguar si el factor(g) "Inteligencia general", (V) "Aptitud Verbal", (K) "Juicio Espacial" y (F) "Habilidad práctica o de Ejecución" podían medirse con precisión a esa edad, o a qué edad podía obtenerse una medida fiable de dichas aptitudes.

Price (1940) administra una batería compuesta por test de referencia de "Inteligencia general" (g); test impresos de juicio espacial y test manipulativos y los de tablero de formas, a una muestra de 85 estudiante de distintas edades.

Después de extraer el factor general, encuentra un factor bipolar, que se aparta de los tests de "g", y que interpreta como "K", ya que la característica común de las pruebas que lo definen está relacionada con figuras y juicio espacial; por lo que deduce que los factores "F" y "K", miden el mismo aspecto de la inteligencia.

Slater (1940, 1941) y, Slater y Bennett (1943) investigan sobre el influjo, de los factores apreciados por los tests espaciales, verbales y no verbales de inteligencia, sobre la edad de los sujetos.

En la investigación de 1940, administra una batería de tests, a una muestra de ochenta y nueve aprendices industriales, de dieciocho años.

Obtiene, además de “g”, un claro factor “K” abarcando a todos los tests espaciales y mecánicos, pero no surge el factor “m” como un componente separado; lo que le lleva a la conclusión de que “juicio espacial”, “K” y “aptitud mecánica”, “m”, eran lo mismo.

En 1941 administra otra batería de tests a una muestra de ochenta y dos muchachos de once años, tratando de verificar los resultados obtenidos por Kelley en América, respecto a que los factores verbal y espacial, podían medirse a los diez años de edad, ya que la confirmación de estos resultados sería muy útil, para afrontar los procesos de orientación y selección, hacia los tres tipos de enseñanza proyectados. Obtiene dos factores, el primero lo interpreta como “g”, pues todas las pruebas saturan en él; el segundo factor, se caracteriza por las elevadas saturaciones que presentan las pruebas verbales, por lo que lo interpreta como “verbal”. Con estos resultados, concluye que si el éxito en la enseñanza secundaria depende, en cierto grado, de la inteligencia general y de la aptitud verbal, ambos tipos de aptitudes pueden medirse con exactitud a la edad de once años. Respecto a la enseñanza técnica, donde el éxito del trabajo parece depender del juicio espacial, éste, es difícil de medir en esa edad.

En 1943 repiten la investigación con dos muestras, una de doscientos once muchachos de once y más años, y, otra, con ciento sesenta y uno de trece y más años; formadas ambas por la misma proporción de varones y mujeres; volviendo a obtener los factores “g” y “v”, lo que le lleva a pensar que los tests de juicio espacial utilizados miden los mismos aspectos que los tests no verbales de “g"; por lo que concluye que este tipo de tests, de juicio espacial, no pueden utilizarse a los once años, e incluso a los trece años, en la orientación y selección de alumnos, pues a estas edades las pruebas de juicio espacial miden inteligencia.

Adcock (1948) reanaliza los datos de Slater y Bennett (1943) ya que le parecía extraño que no surgiera ningún factor distinto de “g” en muestras de once y trece años, lo que atribuye al método de rotación utilizado. Usando el método de Thurstone, múltiple group method, que conduce a un factor general y varios factores de grupo, obtiene una matriz en la que el “juicio espacial”, “K” surge como un factor de grupo, aunque con bajas saturaciones. Señala también, que la controversia Spearman-Thurstone, puede resolverse con el descubrimiento de Thurstone de factores en el dominio de segundo orden.

Emmett (1949) intenta demostrar la existencia del factor espacial a la edad de once años; en el análisis de los datos de Slater obtiene tres factores significativos: “g”, “m” y “K”, en el que alguna de las pruebas no verbales de “g”, aunque no todas, miden el factor “K”. Señala, que en las muestras utilizadas por Slater, los grupos más jóvenes estaban representados en igual proporción por ambos sexos, mientras que el grupo de mayor edad, estaba formado solamente por varones, y las más recientes investigaciones parecían indicar que el factor “K” aparecía menos diferenciado entre las mujeres que entre los varones.

En un estudio realizado por Emmett sobre la muestra de ciento setenta y ocho jóvenes de once y doce años de edad, obtiene que los tests espaciales utilizados, definen un inequívoco factor espacial, en el que los tests espaciales que implican juicios tridimensionales comparados con los bidimensionales, tienen saturaciones más elevadas en “K” que en “g”.

El trabajo de Drew (1947) se orienta en el mismo sentido que el de Slater; trata de medir la aptitud técnico-espacial, a la edad en que se produce el cambio de enseñanza primaria a la secundaria. Los resultados de Slater parecían indicar la imposibilidad de medir el juicio espacial a los once, e incluso a los trece años; y, Alexander y El Koussy, trabajando independientemente habían obtenido los factores “K” y “F” en adolescentes de quince y dieciséis años. Por otra parte, considera que existe un notable solapamiento entre los resultados de lo estudios realizados sobre las aptitudes que se consideran importantes en la formación técnica, aunque no hay ningún tipo de coordinación entre ellos, ya que lo factores "m", "F" y "K", parecen constituir una parte esencial del proceso mental requerido para alcanzar el éxito en el campo técnico.

Con el fin de esclarecer estos resultados, elabora una batería compuesta por tests de referencia de "g", "v" y "F", e incluye un tests de relaciones espaciales y además, considerando que los tests impresos de "K" se limitan al espacio bidimensional y no requieren una manipulación efectiva con objetos concretos tridimensionales, incluye las pruebas manipulativas de Alexander, como medidas del factor "F", así como criterios de una serie de medidas relacionadas con el rendimiento escolar.

Administrada esta, factoriza la matriz de intercorrelaciones por el método centroide y obtiene cinco factores, que le llevan a las siguientes conclusiones:

- Los grupos de edad de once y doce años, ponen en juego una aptitud espacial para el ejercicio de "g" y que, como le ocurrió a Slater, el factor espacial se manifiesta a los dieciséis años, no pudiendo demostrar que los tests espaciales midan juicio espacial a los once años.
- Los tests manipulativos definen un factor, que corresponde al "F" de Alexander, y el test espacial no mide este factor, por lo que el factor espacial es distinto del factor de ejecución "F"; considerando que se verifican los resultados obtenidos por Alexander.
- La entidad psicológica que constituye la "aptitud técnica" son los factores "F" y "K"; siendo la primera vez que se demuestra, mediante el análisis factorial, que dicha aptitud abarca, además de "g", los dos factores de grupo "F" y "K", aunque por otra parte, no muestran una marcada diferencia entre sí, ya que "F" parece exigir un ejercicio mental de orden concreto, con manipulación de materiales, en cambio "K" está asociado con relaciones espaciales de un orden abstracto, motivo por el cual, El Koussy cargaba el acento sobre la imaginación y manipulación abstracta.

En resumen, los factores "g", "F" y "K", son significativos para determinar la "aptitud técnica", concepto más amplio que la "aptitud práctica", por incluir la aprehensión abstracta o imaginativa de las relaciones espaciales. Esta aptitud puede medirse con precisión a partir de los trece años, utilizando una escala manipulativa como la de Alexander1; el test impreso de relaciones espaciales utilizado en el estudio, mide "K" a los dieciséis años, pero no antes, ya que en ese caso miden inteligencia general.

Por tanto puede realizarse la selección y orientación de alumnos para la enseñanza técnica, mediante una escala manipulativa, como medida de la "aptitud técnica", junto con tests de aptitud general y verbal, ya que estos alumnos, además de "aptitud técnica" requieren un mínimo de las otras aptitudes.

Emmett (1949) al reanalizar los datos de Drew (1947), obtiene que los tests manipulativos de Alexander, además de tener saturaciones en "g", aparecen sobre otro factor, con elevadas saturaciones, junto con el test espacial y un test no verbal de "g", interpretando este factor como de "juicio espacial", "K".

---

[1] La Escala de Alexander evalúa la inteligencia teórica-práctica. Esta escala aprecia inteligencia práctica, facilidad de adaptación a distintos ambientes y situaciones, y habilidad en la ejecución. Se basa en la existencia de un factor F, denominado de inteligencia práctica o de "performance", claramente distinto al factor Verbal preponderante en la mayor parte de las pruebas de inteligencia. Utilizada tanto en procesos de selección como en orientación, resulta de gran utilidad para la evaluación de personas con dificultades verbales o desconocimiento del idioma. Está compuesta por tres subpruebas: Passalong (Elevada relación con los test de inteligencia y no implica destreza manual), Cubos de Kohs (Mide desarrollo mental, inteligencia concreta y deterioro mental. También evalúa la expresión analítico-sintética del pensamiento conceptual(factor de estructura espacial)), y Construcción con Cubos. Puede aplicarse a sujetos con deficiencias auditivas.

Vernon (1971) al referirse a estas investigaciones, señala que el error de Drew fue, identificar "F" en función del Passalong, que es la prueba menos fiable de la escala de Alexander y cuando se analiza factorialmente junto con los bloques de Kohs y construcción con cubos, la identidad de "K" y "F" es obvia.

Al reanalizar Vernon los datos obtenidos por Drew con una muestra de setenta y ocho adolescentes, de dieciséis y más años, de la escuela técnica, obtiene tres factores: juicio espacial, "K"; verbal, "v" y, un factor de escolaridad "X" abarcando a todas las calificaciones.

Las investigaciones de Bernyer, en Francia

En Francia, destacan las investigaciones realizadas por Bernyer (1945, 1948, 1949 y 1950) en una línea prácticamente equivalente, conceptual y metodológica, a la inglesa, pero teniendo en cuenta los resultados obtenidos por la escuela americana.

En 1945 publica los resultados obtenidos en una muestra de trescientos veinte aspirantes a piloto, a los cuales administra una batería de tests, de la cual obtuvo veintiocho medidas, que factoriza por el método de Delaporte, que asume la existencia de un factor general, común a todas las pruebas, y una serie de factores de grupo ortogonales. Obtiene, además del factor general, cuatro factores de grupo significativo: uno, relativo a la inteligencia mecánica; otro que abarca las pruebas visuales; un tercero que interpreta como de razonamiento inductivo y; otro que considera como de tipo emocional.

En 1949 publica los resultados obtenidos con una muestra de ciento sesenta adolescentes con edades comprendidas entre trece y diecisiete años, a los que administra una batería compuesta por nueve tests impresos y once manipulativos. El análisis factorial, realizado por el mismo método, indica la presencia de un factor general, seis factores de grupo y seis dobletes.

Interpreta el factor general como una aptitud para captar claramente la naturaleza de la tarea en su conjunto y movilizar las destrezas más particulares implicadas en la tarea. Los factores de grupo los interpreta como: inducción, deductivo, aptitud para visualizar el espacio, un cuarto que interpreta como "inteligencia práctica" aplicada a problemas mecánicos en los cuales el sujeto recurre a experiencias adquiridas; un quinto, de tipo manual y el sexto de "agilidad manual" semejante al de Harrell.

En 1950 intenta confirmar el factor de tipo práctico obtenido en la investigación anterior, pero no lo logra.

Escuela Americana

TRABAJOS EN LAS FUERZAS AÉREAS (A.A.F.).

Las investigaciones Americanas en relación con la inteligencia técnica, a primeros de los años de la década de los cuarenta, durante la segunda guerra mundial, se centra sobre los problemas planteados por la selección y clasificación de personal para desempeñar tareas técnicas y mecánicas de tipo bélico. En este campo de investigación destacan los trabajos realizados en las fuerzas áreas (A.A.F.) por un grupo de psicólogos encabezados por Flanagan, intentando descubrir las aptitudes más importantes para el eficaz desempeño de los puestos de trabajo relacionados con las tripulaciones aéreas. Terminada la contienda, en 1947 publican todos los

trabajos terminados, los interrumpidos y los resultados obtenidos hasta entonces, en los diecinueve volúmenes del "Army Air Aviation Psychology Program Research Report".
El estudio psicológico de todas las fuentes de información aprovechables, permitió confeccionar una lista con los veinte rasgos que fueron considerados como más importantes para el éxito de los pilotos; estos veinte rasgos fueron agrupados en cuatro categorías: Intelectual (I), Perceptivo (P), Temperamental (T) y Psicomotor (M).

El principal método de investigación utilizado fue el análisis de las correlaciones, en la validación de tests y baterías, la regresión simple y múltiple, y el análisis factorial para definir la naturaleza de la contribución de los tests en el pronóstico del criterio. En el report nº 5 señalan que un total de veinticinco mil ochocientos cuarenta y cuatro sujetos se distribuyeron entre los dieciocho análisis factoriales realizados, dirigidos por Guilford; el análisis con menor número de sujetos se basa en una muestra de ciento setenta y el de mayor en ocho mil ciento cincuenta y ocho, basándose en ciento ocho tests y, obteniendo veintisiete factores (Pascual, 1975).

Entre los factores más conocidos surgían, en casi todos los análisis, los de : comprensión verbal, facilidad numérica, rapidez perceptiva y memoria repetitiva; entre los nuevos factores descubiertos se encontraban los de: visualización, coordinación psicomotora y experiencia mecánica y; entre los sugeridos o insinuados como posibles hipótesis de trabajo los de: razonamiento, planeamiento, juicio, memoria, estimación de longitudes y los de rapidez y precisión psicomotora.

Estas cifras y datos nos sugieren lo extenso de las investigaciones realizadas en el A.A.F., por lo que solo vamos a limitarnos a reseñar los trabajos de Guilford (1947) sobre los factores del área perceptiva, los psicomotores y las investigaciones en el campo de la "aptitud mecánica".

INVESTIGACIONES SOBRE LOS FACTORES "PERCEPTIVOS", "PSICOMOTORES" Y DE "APTITUD MECÁNICA".

FACTORES DEL ÁREA PERCEPTIVA

Rapidez perceptiva; aptitud para comparar rápidamente formas visuales y reconocer con precisión las semejanzas y diferencias entre las figuras y los detalles. Semejante al factor "P" de Thurstone.

Estimación de longitudes; exige la comparación de líneas o distancias entre puntos.

Factores espaciales; el análisis de la matriz factorial de los primarios de la batería de Thurstone con el método de rotación de Zimmerman, permite desdoblar el primario "S", que Thurstone había interpretado como la aptitud para imaginar y reconocer objetos y estructuras que se mueven en el espacio, en tres factores:

> Primer factor espacial; sugiere una aptitud para percibir el orden espacial o las relaciones entre los objetos; una orientación espacial en la que parece importante la referencia al propio cuerpo, cuando se considera como el origen de coordenadas.
>
> Segundo factor espacial; el aspecto común a los tests que lo definían; manos, banderas, figuras y tarjetas de Thurstone, que a su vez tenían saturaciones en el primer factor; parecía consistir en una discriminación en la apreciación de las manos derecha - izquierda, sugiriendo la posibilidad de que pueda abarcar la imaginación cinestésica. French (1951),

en su estudio de factores, no encuentra clara la naturaleza de este factor y lo refiere al también dudoso de orientación espacial.

Factor de visualización espacial; fue considerado como una función dinámica requerida en tareas que implican el movimiento de máquinas, la transformación de objetos y los cambios de posición en el espacio. La solución de los problemas presentados exigen del sujeto una manipulación mental para mover, girar o rotar un objeto u objetos, y reconocer su nueva apariencia o posición después de realizar la manipulación descrita.

FACTORES DEL ÁREA PERCEPTIVA

Factor psicomotor I o "coordinación psicomotora"; común a los tests psicomotores y surge en todos los análisis que incluían dos o más tests de este tipo; abarcando tanto pruebas que exigen pequeños ajustes musculares como los que requieren el uso de los grandes músculos, movimiento de brazos, piernas y tronco; obteniéndose las mejores medidas del factor con pruebas que reclaman movimientos de alcance moderado.

Factor Psicomotor II, tentativamente denominado de "precisión psicomotora", ya que parecía requerir movimientos precisos y realizados con rapidez, a presión del tiempo. French (1951) lo identificó con "destreza digital", ya que exige la manipulación rápida de objetos con ambas manos; siendo distinto de la "destreza manual" por que no incluye los movimientos de brazo y de "habilidad de puntería" (aiming) al no precisar la coordinación ojo - mano.

Factor psicomotor III, tentativamente denominado como "rapidez psicomotora" ya que surge en pruebas cuyo unido elemento común es la rapidez. French (1951) lo interpreta como de "puntería" ya que considera que representa la habilidad para realizar con gran rapidez movimientos que requieren la coordinación ojo - mano, asemejando la operación de marcar una hoja de respuestas con la tarea requerida por los tests de trazado, marcado y punteado de McQuarrie.

"APTITUD MECÁNICA"

Las investigaciones anteriores a la guerra en el campo técnico y mecánico, parecían indicar que el éxito en los trabajos referentes a la comprensión, utilización y manejo de máquinas y mecanismos, exigía ciertas aptitudes especiales y que, además, algunas tareas relacionadas con las máquinas no se limitaban exclusivamente a los trabajos que tradicionalmente se consideraban como mecánicos.

Realizados los análisis de tareas de los puestos de tripulaciones aéreas, elaboran una batería experimental que constaba de siete pruebas de mecánica, comprensión e información, tres de estimación de longitudes, dos de comprensión y análisis de patrones, varios tests impresos y el test de coordinación compleja. Administrada la batería a una muestra de ciento cincuenta y tres alumnos que seguían un curso previo para su posterior clasificación entre los puestos de tripulaciones aéreas; factorizan la matriz de correlaciones resultantes, obteniendo siete factores:

Experiencia mecánica, definido por los siete tests de mecánica.

Rapidez perceptiva, definido por tests que requieren del sujeto una gran rapidez para continuar, imaginativamente, los movimientos de las partes de las máquinas.

Verbal; sobre el que aparecen con bajas saturaciones los tests de comprensión mecánica y razonamiento aritmético.

Estimación de longitudes; definido por pruebas perceptivas, basadas en estimaciones de tipo cualitativo, que presentaban ciertos movimientos de distracción.

Visualización; definido como de manipulación visual o visualización, pues los tests con saturaciones más elevadas en él, muestran objetos que deben imaginarse en movimiento, o que han realizado ese movimiento con anterioridad o bien que los objetos han sufrido una transformación.

Relaciones espaciales; donde el aspecto más significativo y sobresaliente del factor es el orden espacial de los ingenios mecánicos que constituían los elementos de los tests.

Razonamiento general, definido por tests de razonamiento aritmético.

Las conclusiones del estudio indican que la principal aportación del análisis era, por una parte, la mejor, más clara y precisa definición de los test de mecánica, y por otra, el descubrimiento de un factor común y exclusivo de todas estas pruebas, el factor de “información o experiencia mecánica”; las pruebas de conocimientos mecánicos eran las medidas más puras de este factor.

La “aptitud mecánica” considerada comúnmente como una aptitud unitaria, se encuentra constituida por un complejo de varias aptitudes. En función de la clase de tests y de las diversas clases de trabajo mecánico, sugiere Guilford que, además de la experiencia mecánica, los otros aspectos importantes de esta compleja aptitud son los factores de “visualización”, “relaciones espaciales”, “rapidez perceptiva” y, en algunos casos, el factor de estimación de longitudes.

EL TRABAJO DE DEGAN, SOBRE LA “APTITUD MECÁNICA”

Degan (1950) publica un estudio realizado de los tests de “aptitud mecánica” del A.A.F. en el laboratorio de psicometría de la Universidad de Chicago. Partiendo de la matriz centroide del A.A.F. y siguiendo el criterio de estructura simple, obtiene una matriz oblicua que le permite estudiar los factores primarios en el dominio del segundo orden.

En el análisis de primer orden obtiene siete factores, que interpreta de forma parecida al A.A.F.

Información o experiencia mecánica; definido por tests que reclaman conocimientos sobre la naturaleza y función de las herramientas, así como el significado del vocabulario técnico y las operaciones mecánicas asociadas con esas palabras. Opina, que probablemente, pueda considerarse este factor, en un sentido rescindido, como un criterio parcial, de una medida de la “aptitud mecánica”.

Rapidez perceptiva; inicialmente interpretado como de “rapidez de identificación”, y en algunas ocasiones como de “orientación espacial”, aunque en la mayoría de las ocasiones se interpretó como “rapidez perceptiva”, ya que su naturaleza es perceptiva y de rapidez, aunque algo más complejo, que el denominado “rapidez perceptiva” por Thurstone.

Verbal - escolar; interpretado por el A.A.F. como “verbal”, pero Degan, lo denomina así, ya que considera que además, depende de la formación escolar de los sujetos.

Flexibilidad o fuerza de clausura; en el A.A.F. fue interpretado como estimación de longitudes, pero Degan lo interpreta así, al considerar que además de estimar longitudes, la mayoría de los elementos muestran un campo perceptivo que distrae al sujeto y, a veces, llega a producir "ilusiones".

Visualización dinámica; identificada como visualización por el A.A.F.

Cibernético; inicialmente considerado por el A.A.F. como el componente intelectual del aparato de coordinación compleja, pero Degan señala que, la característica más general de todos los tests con saturaciones significativas en el, es la facilidad para controlar o determinar la dirección del movimiento respecto a la orientación del sujeto.

Razonamiento; todas las pruebas requieren del sujeto procesos complejos de razonamiento.

En el dominio del segundo orden, obtiene Degan tres factores. Un factor, que denomina "K", formado por los de rapidez perceptiva y verbal escolar, que viene a representar un factor complejo dentro del arara de la aptitud de lectura. Otro, que denomina, "L", definido por los de razonamiento, flexibilidad de clausura y, en cierta medida, por el cibernético, que probablemente, podría sugerir la noegénesis de Spearman. Y un tercer factor "J", definido por los primarios de experiencia mecánica, espacial dinámico o visualización y cibernético, que viene a representar aquellos elementos, que se combinan para producir destrezas tales como la "aptitud mecánica"; destacando la estrecha relación entre esta y los factores espaciales, así como su relación con el razonamiento, y su independencia del verbal educativo.

PROGRAMA DE INVESTIGACIÓN "USES"

En 1945, la dirección de análisis de ocupaciones, programa de investigación de oficios del servicio de empleo de los Estados Unidos, "USES", publica un resumen de los trabajos realizados sobre la batería de tests factoriales USES.

Los trabajos, iniciados en 1934, dirigidos a la orientación escolar y profesional, influyeron sobre el concepto de aptitudes "especiales" y sobre la arbitraria diferenciación entre tests de inteligencia, aptitudes y conocimientos, considerados como categoría mutuamente excluyentes. Tanto los autores del proyecto Minnesota como los psicólogos de la época, suponían que las aptitudes "especiales", a semejanza de las intelectuales, eran disposiciones innatas independientes del aprendizaje y la experiencia. La evolución de este concepto les lleva a considerar que las respuestas de los sujetos en los tests, dependían, por supuesto, de las disposiciones heredadas, pero también del influjo particular que las diversas oportunidades de aprendizaje proporcionan a lo largo del desarrollo y maduración del individuo.

Basándose en los resultados obtenidos en otras investigaciones, proyecto Minnesota, Harrell, Thurstone y las del A.A.F., entre otras, tratan de encontrar los factores más útiles y significativos en la práctica de la orientación profesional.

Siguiendo este criterio seleccionan cincuenta y nueve tests, cincuenta impresos y nueve manipulativos; analizan las tareas de más de quinientas profesiones, y teniendo en cuenta los resultados de los análisis factoriales, las agrupan en veintitrés subconjuntos afines, para confeccionar los correspondientes profesiogramas a cada familia profesional y así; elaborar las baterías de aptitudes generales, en lugar de componer baterías específicas para cada profesión.

Distribuyen los cincuenta y nueve tests seleccionados, en baterías de quince a veintinueve pruebas, que administran a dos mil ciento cincuenta y seis sujetos, divididos en nueve grupos experimentales; con edades comprendidas entre diecisiete y treinta y nueve años.

En los nueve análisis factoriales se utilizan el método centroide y el criterio de estructura simple para la rotación de ejes, obteniendo estructuras aproximadamente ortogonales. En cuatro de los análisis obtienen diez factores; en uno, nueve; en dos ocho y; en los dos restantes, siete.

Algunos de los factores fueron identificados como bien conocidos; verbal; numérico; espacial, abarcando todos los tests bi y tridimensionales; experiencia mecánica y; escolaridad.

Entre los restantes, dos de tipo perceptivo, uno con figuras de tipo espacial y mecánico, "P" y otro, con comparaciones de nombres y números, "a"; un factor de puntería, "Aiming"; otro de velocidad, muy relacionado con el de puntería; dos de destreza, uno digital "F" y otro manual "M" y, por último, uno que designan "lógico" que surge en dos estudios y que abarca, pruebas de analogía y de instrucciones verbales, requiriendo la solución de problemas mediante procesos de razonamiento formal.

Basandose en los resultados, construyen una batería con diez tests impresos y dos manipulativos, para evaluar los nueve factores siguientes: inteligencia, verbal, aptitud mecánica, espacial, percepción de formas, y percepción burocrática, coordinación motora y destrezas digital y manual.

TRABAJOS DE THURSTONE SOBRE EL "CAMPO PERCEPTIVO"

Thurstone (1944) al realizar una investigación para explorar el campo perceptivo y sus posibles relaciones con la dinámica de la personalidad, para la cual utiliza material perceptivo y así, poder estudiar los efectos producidos por "Gestalten" visuales, presenta a los sujetos un material ambiguo, ante el cual tiene que emitir un juicio comprensivo, claro y definido, lo que le permitiría estudiar las diferencias individuales.

Con tal fin, selecciona cuarenta pruebas que tratan de representar una gran variedad de fenómenos perceptivos, la mayoría consistentes en aparatos de aplicación individual, que administra a ciento setenta sujetos voluntarios. Al factorizar la matriz de intercorrelaciones obtiene once factores, de los cuales, el último lo considera residual y dos, como dobletes; otro lo interpreta como "general intelectivo" ya que venía definido por los primarios: verbal, razonamiento y numérico; tres parecían representar algún tipo de velocidad o rapidez en las respuestas de los sujetos: tiempo de reacción, velocidad de percepción y rapidez de juicio; dos, se referían al material específico utilizado en el experimento: ilusiones con dibujos geométricos y oscilación o ritmo de fluctuación de figuras reversibles vistas en perspectiva y; por último, los dos restantes, se referían a la clausura, el factor de rapidez y fuerza de clausura y el factor "Gestalts cambiantes".

El primero de estos, "rapidez y fuerza de clausura" representa la aptitud para formar una unidad perceptiva en un campo desorganizado a pesar de los elementos de distracción.

El segundo, "Gestalts cambiantes" representa la aptitud para manipular dos configuraciones simultánea o sucesivamente, o también, la aptitud para anular una configuración y establecer otra. Todas las pruebas que lo definían, tenían en común el requerir del sujeto un esfuerzo para liberarse de la sujeción que ejercen formas previas y usuales, romperlas y sustituirlas por otras estructuras. Vernon (1971) sugiere que esta función corresponde al factor "g".

El descubrimiento de estos factores de clausura motivo una serie de investigaciones, tanto de Thurstone como de sus seguidores, tratando de aclarar su naturaleza y sus posibles relaciones con los factores lógicos y aspectos prácticos de la personalidad.

Un intento de aproximación entre ambas escuelas

El Koussy (1948) durante el Congreso Internacional de Psicología celebrado en Edimburgo, señala que en un análisis de los tests espaciales de Thurstone realizado por el método de las diferencias tetrádicas, sugiere la posibilidad de que los factores “K” y “S” sean un mismo factor, ya que este se mide mediante pruebas que reclaman la aptitud para visualizar y manipular mentalmente el material espacial y, piensa, que los tests espaciales tridimensionales proporcionan mejores medidas de esta aptitud que los bidimensionales, proponiendo como hipótesis de trabajo averiguar las relaciones existentes entre los factores “m”, “K” y “F”.

Por los trabajos aquí reseñados y los resultados de ulteriores investigaciones, los factores “K” de El Koussy y “S” de Thurstone, se refieren a la misma aptitud de visualización espacial y; los factores “m” de Cox y “F” de Alexander, dependen en gran medida, de la referida “aptitud de visualización” y, además, vienen afectados por otros aspectos comunes, cognoscitivos y no cognoscitivos y, de la personalidad de los sujetos, como por ejemplo, el factor de “voluntad de éxito”, “X”, de Alexander.

Bibliografía

ADCOCK, C. (1.948). “A Reanalysis of Slater's Spatial Judgment Research”. Occup. Psychol., 22: 213-216.

BERNYER, G. (1.945). “Un essai d'analyse factorielle des aptitudes”. Ann. Psychol., 41-42: 202-226.

BERNYER, G. (l948). “Distribution des facteurs psychologiques dans une population”. Ann. Psychol., 45-46: 16-28.

BERNYER, G. (1.949). “Analyse factorielle de tests d`aptitudes techniques”. Ann. Psychol., 47-48: 198-211.

BERNYER, G. (1.950). “Recherches sur quelques tests d'aptitude apliqués aux apprentis d'une école de mécanique”. Ann. Psychol., 49: 159-171.

DEGAN, J. W. (1.950). *A re-analisys of the Army Air Force battery of mechanical test. Psychometric Laboratory*. Univ. Chicago. núm. 58.

DREW, L. J. (1.947). “An investigation into the measurement of technical ability”. Occupational Psychology 21: 34-48.

EL KOUSSY, A. A. H. (1.956). “Les directions de recherche dans le domaine des aptitudes spatiales”. En *L`analyse factorielle et ses applications*. París, C.N. Rech. Scientif., 327-351.

EMMETT, W. G. (1.949). “Evidence of a space factor at 11 plus and earlier”. British Journal of Psychology, (Statistics Section). 2: 3-16.

FRENCH, J. W. (1.951). *The description of aptitude and achievement test in terms of rotated factors*. Psychometric Monograph, nº 5.

GUILFORD, J. P. and J.L. LACEY. (1.947). *Printed Classification Tests*, A.A.F. Aviation Psychological Progress Research Report, nº 5, Washington, D.C. U.S. Government Printing Office.

PASCUAL QUINTANA, M. (1.975). "Estructura y dimensiones de la Aptitud de Vuelo". Rev. Ps. Gral. y Apld., 30, 133 y 134: 287-332.

PRICE, E.J.J. (1.940). "The nature of the practical factor, (F)". British Journal of Psycology 30: 341-51.

SLATER, P. (1.940). "Some group tests of spatial judgement or practical ability". Occupational Psychology 14: 40-55.

SLATER, P. (1.941). "Test for selecting secondary and technical school children". Occupational Psychology 15: 10.

SLATER, P. Y BENNETT, E. (1.943). "The Development of Spatial Judgement and its Relation to Some Educational Problems". Occupational Psychology 17: 139-155.

THURSTONE, L.L. (1.944). *A factorial study of perception*, Chicago, University of Chicago Press.

VERNON, P.E. (1.947 a). "Research on Personnel Selection in the Royal Navy and the British Army". Amer. Psychologist, 2: 35-51.

VERNON, P.E. (1.947 b). "Psychological Tests in the Royal Navy, Army and A.T.S.". Occup. Psychol., 21: 53-74.

VERNON, P. E. (1.971). *The Structure of Human Abilities*. Methuen & Co. London, Nueva impresión de la 2a edición de 1.961.

VERNON, P.E. & PARRY, J.B. (1.949). *Personnel Selection in the British Forces*. London: University of London Press. Pp. 324.

Capítulo 33

Consideraciones sobre la medida de la concentración por Gini

ROMERO GARCÍA, JOSÉ ENRIQUE
GAMERO ROJAS, JAVIER
Universidad de Sevilla

Introducción

A finales del siglo XIX y principios del XX en diversos países europeos (y en Estados Unidos) se estaban desarrollando rápidamente las primeras técnicas de la Estadística Descriptiva. Hasta cierto punto en este desarrollo hubo una "especialización" en diferentes países, así por ejemplo en Inglaterra se estaba desarrollando los campos aplicados a la psicología, biología, ciencias sociales, etc.; en Francia y en Rusia, por otra parte, existían considerables aportaciones con un enfoque más teórico y en Italia hubo una escuela de estadísticos aplicados al campo de la economía y concretamente al campo de las cantidades "macro" de los Estados. Ni que decir tiene que esta descripción esquemática no tiene vocación de ser exhaustiva en absoluto.

En este contexto es donde aparecen figuras como las de Pareto, Benini, Gini,… que estudian, entre otras cosas, la distribución y concentración de magnitudes relevantes para la macroeconomía (rentas, riqueza de individuos, tamaño de familias, créditos,…). Corrado Gini estudió ampliamente estos puntos y, como algunos otros estadísticos italianos, lo hizo fundamentalmente desde un punto de vista aplicado y, casi podríamos decir empírico. Gini, en el fondo, mantuvo una cierta desconfianza sobre el papel de las Matemáticas en la aplicación de la Estadística. Reproducimos a continuación algunas frases de la edición inglesa The Contributions of Italy to Modern Statistical Methods
por Corrado Gini.

"for Statistics they [Mathematics] are no more than a means".

"This does not mean, however, that it is always convenient to use this method[Mathematics] ; the statistical material at our disposal may be too rough to allow of the application of exquisite methods, or for the purpose of the research it may appear useless to seek precision beyond a certain limit"

"His motto [of a statistician], instead of being represented by the words "Statistics with Mathematics," should be represented by "Statistics with the least mathematical means possible."

"I mean to say that there are wide fields in which... the statistical analysis does not need any instrutnents of calculation higher than simple arithmetic, and where it would be not only useless but harmful to have recourse to refined mathematical methods which, while they would not increase appreciably the fruitfulness of the research, would reduce the number of readers".

"Mathematics [they are not], however, sufficient to ensure success, which above all depends upon a kind of statistical intuition helped by wide experience".

Algunas de estas frases son muy interesantes en más de un sentido, pero no es nuestro propósito aquí el extendernos sobre este asunto, sino sólo dar una muestra de la mentalidad de Gini sobre la investigación analítica en la estadística.

El índice δ de Gini

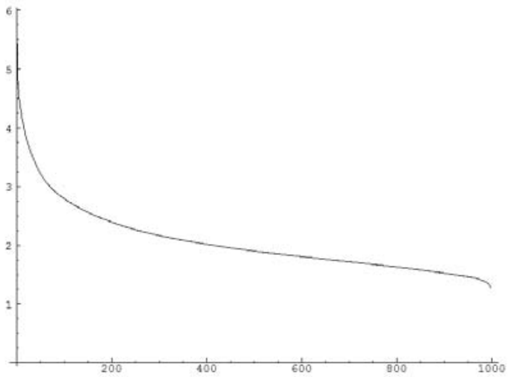
Siguiendo las ideas de Pareto y considerando las relaciones empíricas que observó con los datos disponibles de reparto de rentas en diferentes territorios y periodos de tiempo, Gini vio que, aproximadamente, se verificaba la siguiente relación:

$$1-P_i = (1-Q_i)^{\delta} \qquad (1)$$

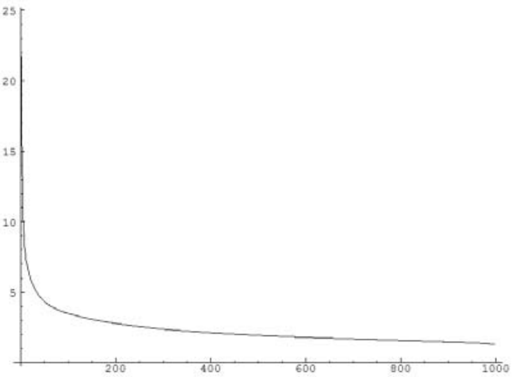
En donde, usando la notación más establecida, Pi es la frecuencia relativa acumulada en el individuo i-ésimo y Qi es la cantidad relativa acumulada en ese mismo individuo i-ésimo (supuesto que los individuos están ordenados de menor a mayor renta).

Si esta relación se mantuviese cierta para todo “i” con un mismo δ**,** naturalmente este valor podría ser un indicativo de concentración de rentas en individuos. Con esta suposición, Gini calcula su coeficiente de concentración δ como un “valor medio” entre todos los “deltas” δi que se irán obteniendo para los diferentes individuos o clases “i”. Entendemos que este hallar el "valor medio" es un "ajuste" de la distribución observada al modelo (1)

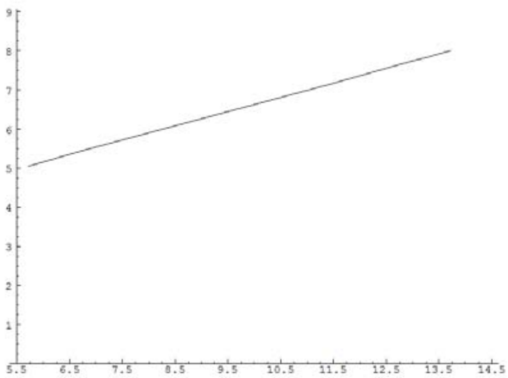
Este enfoque presenta el lógico problema de que no es adecuado cuando la relación (1) no es correcta, en el sentido de que para cada “i” hubiese valores “delta” claramente distintos, y peor aún, cuando los valores delta cambian según un patrón al crecer “i”. En realidad, sólo si la distribución de las rentas es de tipo Pareto, los valores delta serían teóricamente iguales. Con distribuciones tipo LogNormal o Gamma, por ejemplo, los deltas disminuyen sistemáticamente según crece “i”. En otros casos, como el propio Gini observará en la distribución del número de hijos por familia, los deltas crecen sistemáticamente. A continuación incluimos unos gráficos con los valores δi para diferentes supuestos distribucionales y ejemplos.



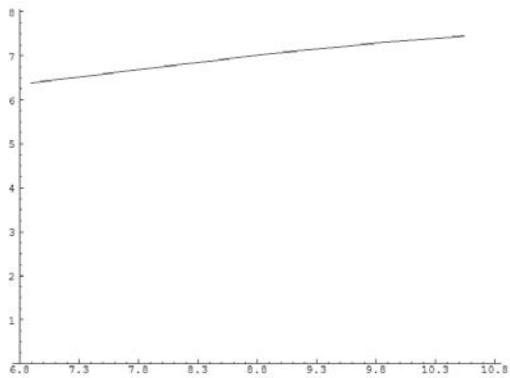
Distribución LN(0'5,1)



Distribución g(4,1)



Ejemplo de Gini pág.16



Ejemplo de Gini pág. 18

Los dos últimos gráficos corresponden a los ejemplos analizados por Gini en su tratado *Indici di concentrazione e di dependenza* en las páginas 16 y 18. Puede observarse que en ninguno de los cuatro casos el valor δ i es estable y por tanto no puede hablarse de un único δ que represente a la distribución.

En el caso de una distribución tipo Pareto el gráfico hubiese sido aproximadamente horizontal (con las posibles fluctuaciones muestrales, en su caso), por tanto la no horizontalidad viene a reflejar la diferencia de la distribución analizada con la Pareto. Así, veríamos que el segundo ejemplo de Gini es el más próximo a un caso paretiano.

Por tanto, en un amplio abanico de posibilidades realistas, no se puede hablar de "un" valor de δ que ligue Pi y Qi, y por consiguiente el propósito de este índice queda desvirtuado, excepto en distribuciones paretianas.

Aunque apoyado en la distribución de Pareto, este índice δ tiene, sin embargo, a ojos de Gini, varias ventajas sobre el índice δ de Pareto. En lo que sigue exponemos el texto original de Gini y posteriormente un comentario nuestro encabezado por una "síntesis" entrecomillada.

"Ci pare che il nostro indice di concentrazione δ presenti alcuni vantaggi di fronte all`indice di distribuzione α del Pareto.

> L`indice α è atto a ritrarre la distribuzione dei redditi solo al di sopra di un dato limite. Non si misura, dunque, mediante esso, la disuguagianza di distribuzione di tutta la ricchezza, ma solo di una sua parte.
>
> L`indice δ, al contrario, si adatta ugualmente bebe alle seriación di redditi complete (che dànno cioè tutti i redditi da 0 all`insù) che alle incomplete (che dànno solo i redditi al di sopra di un dato limite);

"El índice δ sólo se puede aplicar a rentas por encima de cierto punto (debido a las características de la distribución de Pareto), mientras que δ puede aplicarse a la distribución completa de la renta". Este comentario es criticable en el sentido de lo que hemos señalado anteriormente: δ es plenamente satisfactorio si la distribución es paretiana, pero si lo es, aparece el mismo problema de límite inferior de rentas que caracteriza a esta distribución. Es decir, la justificación plena de δ reside en la propia distribución de tipo Pareto.

(a) L`indice δ ha un campo di validità più esteso di quello dell`indice α.

"El *índice α se aplica en un caso más restringido que el índice δ*". En realidad Gini aplica δ a casos donde no es del todo satisfactorio (cuando los deltas no son constantes para todo i) y eso sería equivalente a aplicar α a distribuciones que no son exactamente paretianas. Como hemos apuntado anteriormente, en el fondo el procedimiento de cálculo "*promedio" de Gini para δ es similar a "ajustar"* α a distribuciones que no sean tipo Pareto.

(b) L`indice δ è molto più sensibile sell`indice α alle differenze di distribuzione dei redditi. Notevoli diffrerenze di distribuzione dei redditi rimangono appena avvertite dai valori di α, specialmente se i valori di α sono bassi. Di qui era sorta presso molti l`idea, come vedemmo infondata, che la distribuzione della ricchezza fosse pressochè uguale in tutti i paesi e in tutti i tempi ;

"El índice δ es más sensible que α en la medición de la concentración". Es decir, se afirma que $\left|\delta'(\alpha)\right| > 1$. Es interesante observar que δ y α son "autoinversos", o sea que $\delta(\alpha)$ es funcionalmente igual a $\alpha(\delta)$ (o, lo que es lo mismo, que esas funciones son simétricas respectos a la primera bisectriz). Por tanto, en sí, α y δ, globalmente, son igualmente sensibles, pero, y en esto Gini está en lo cierto, en el tramo habitualmente observado (δ > 2), δ es más "sensible" que α.

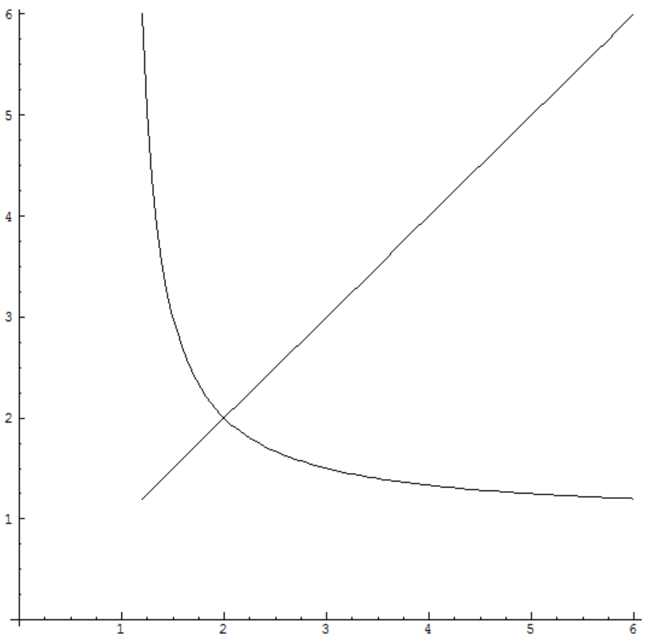


Gráfico de d(a) y a(d)

Sobre este mismo punto, conviene señalar que Gini encuentra una refutación de la "ley" que Pareto creyó encontrar sobre la distribución de la riqueza. Pareto observó que la distribución de la riqueza en diferentes territorios y épocas tendía a seguir una distribución paretiana con un parámetro α similar; a ese tal valor α le quiso dar un significado de "ley" universal.

Gini refuta eso señalando que los valores "similares" de α corresponden a valores bastante distintos de α debido a la relación funcional antes señalada entre ambos parámetros. A su vez, valores algo distintos de δ implican concentraciones muy distintas. Citamos dos párrafos de estadístico italiano al respecto:

"Il Pareto, infatti, tratando le seriación dei redditi con metodo matematico, giunse alla conclusione que la distribuzione dei redditi globali è pressochè identica per tutti gli Stati e per tutti i tempi............

"Notiamo però come, a differenze relativamente piccole nel valore dell`indice di concentrazione, corrispondano grande differenze nella distribuzione della ricchezza.
Se è δ=2, 1/m del reddito tassato sarà posseduto da $1/m^2$ dei redditeri censito; per esempio ¼ dei censito possederà ½ del reddito; 1/9 ne possederà 1/3, 1/16 ne possederà ¼.
Se è invece δ=3, 1/m del reddito tassato sarà posseduto da $1/m^3$ dei redditieri censito; e cosi ½ del reddito si troverà in mano non più di ¼, ma di 1/8 di censito; 1/3 del reddito in mano non di 1/9, ma di 1/27 dei censito; ¼ del reddito in mano non di 1/16, ma di 1/64 dei cenciti."

Proseguimos con la exposición del resto de puntos sobre la comparativa de Gini entre su coeficiente y el de Pareto:

(c) L`influenza perturbatrice dell`evasione si fa sentire, come abbiamo mostrato, meno sull`indice δ che sull`indice α.

(d) L`indice δ ha un significado preciso. Esso indica l`exponente al quale bisogna innalzare una frazione dei redditi accertati per ottenere la frazione dei censito che li possiede. Quando δ crece, aumenta quinde la concentrazione dei redditi; nessun equivoco è possibile."

"El índice δ tiene una interpretación sencilla e intuitiva". Pero α, añadimos nosotros, por supuesto también la tiene: nos da precisamente la distribución y los cuantiles de la renta en forma sencilla.

A través de estos breves comentarios a las apreciaciones de Gini sobre su propuesta de índice en comparación con la medida de concentración de Pareto, podemos observar que no está fundamentada una significativa ventaja de δ sobre α. Aquí conviene traer a colación lo señalado en la sección 1 sobre la relativa desconfianza del estadístico veneciano respecto a la matemática teórica. Posiblemente si Gini hubiese profundizado en el análisis matemático de su coeficiente δ, se hubiese convencido de que no aportaba ventajas significativas y quizás, entrando ahora en lo hipotético, hubiese acelerado su adhesión definitiva al coeficiente que lleva su nombre.

Diferencias entre los conceptos y mediciones de la concentración entre Gini y Pareto

Pareto (1848-1923), aunque nacido en París y de madre francesa, hizo sus estudios y formación en Italia y puede considerarse perteneciente culturalmente a este país. Pareto fue de una generación anterior a Gini (36 años mayor) y es natural que éste tuviese como referencia relativa el trabajo de aquél. En Indici di concentrazione e di dependenza Gini señala la disparidad entre la idea de "concentración" de Pareto y la suya.

"Altrettanto non si può dire dell`indice α, sulla cui interpretazione vi fu e vi è dissenso fondamentale. Mentre il Pareto infatti ritiene che il crescere di α indichi aumento di disiguaglianza nella distribuzione, il Benini ritiene al contrario che esso indichi diminuzione di disuguaglianza. È da avvertire che il dissenso dipende dal diverso significato che i due illustri statistici dànno alle espressioni : maggiore o minore disuguaglianza della distribuzione. Poichè, per il Pareto, la disuguaglianza aumenta quando diminuisce la percentuale dei censiti con reddito superiore ad x, mentre, per il Benini, im tal caso, la disuguaglianza diminuisce.
Vi è certamente in tutte le definizioni qualche cosa di arbitrario. Ma dobbiamo guardarci, a scanso di equivoci, dall`atribuiré ai termini un significado in contrasto con quello che viene loro comunemente attribuito.

Prendendo le parole nel loro significado etimologico e corrente, dobbiamo dire che la concentrazione della riccheza aumenta e la sua disuguaglianza si fa più forte quando diminuisce la frazione dei censito al di sopra di un dato reddito che possiedono una data parte dei redditi accertati, o viceversa quando aumenta la parte dei redditi accertati posseduta da una data frazione di censito al di sopra di un dato reddito. Ora, in tal caso, aumenta il valore di δ e diminuisce corrispondentemente, in teoria, il valore di $\alpha = \frac{\delta}{\delta - 1}$.
È necesario dunque concluyere che l`interpretazione del Benini, secondo il quale il diminuire di α indica un aumento di disuguaglianza nella distribuzione, ricorsponde

meglio che quella del Pareto al sgnificato che comunemente si attribuisce alla espressione <disuguaglianza nella distribuzione>.

Nei fatto però non avviene che ad un aumneto di δ corresponda una diminuzione di α. Per la mancanza di coincidenza fra δ e $\frac{\alpha}{\alpha - 1}$*, può darsi cge al creceré di δcorrisponda un aumneto di α. In Hamburgo, per essempio, il valore di δ è aumentato dal 1883 al 1895, mentre è aumentato pure il valore di α.”*

Es interesante señalar que en la visión de Pareto hay mucha concentración (en rentas) cuando una gran cantidad de personas tienen rentas pequeñas y un reducido número de personas tienen rentas mucho más mayores. Aquí, el economista italiano interpreta la concentración como "concentración de rentas parecidas" (dispersión relativa pequeña). En términos "modernos", siguiendo a Gini, esa situación se consideraría como "concentración pequeña" (dispersión relativa pequeña). Es decir, en términos de Gini "concentración" y dispersión (relativa) son conceptos paralelos, no opuestos,

Conclusiones

En sus trabajos sobre medición de la concentración distribucional de diversas magnitudes (rentas, riqueza, número de hijos,...) Gini fue elaborando a lo largo de los años una serie de coeficientes para determinar numéricamente tal cantidad. En estos sucesivos coeficientes podemos ver antecesores y "parientes" de los índices hoy corrientemente usados. Nos hemos centrado en este trabajo en el índice δ, que está especialmente relacionado con los trabajos de su antecesor Pareto e indirectamente relacionado con lo que conocemos como Curva de Lorenz. Gini se refiere al economista y matemático italo-francés en varias ocasiones para comparar sus descubrimientos y progresos en el campo. Así asevera, en su momento, que su coeficiente δ es superior al coeficiente α de Pareto y lo utiliza para refutar la "ley de Pareto" de igualdad de concentración entre los Estados. De alguna forma, también polemiza con las ideas de su antecesor en lo referente a la idea misma de "concentración" razonando sobre el significado de α y δ, como hemos observado.

Sin embargo, hemos podido señalar que el coeficiente δ no poseía todas las virtudes que Corrado Gini creía ver en él y, sólo hipotéticamente, hemos puesto en relación este exceso de confianza en la relativa displicencia del italiano hacia el método matemático.

Aunque su estudio se sale de los límites impuestos a este pequeño trabajo, cabe señalar que Gini, consciente de algunas limitaciones de su coeficiente δ, propuso simultáneamente otras variantes para diferentes aplicaciones. De esta manera nació el coeficiente τ, o el "coeficiente complejo" (x, ε). Todo ello formó parte del proceso que llevó a Gini y a otros a construir las medidas de concentración actualmente en uso (y que siguen en proceso de evolución).

Bibliografía

GINI, C. (1913), "Sulla misura della concentrazione e della variabilita dei caratteri", Atti del R. Istituto Veneto di Scienze, Lettere ed Arti.

GINI, C. (1922), "Indici di concentrazione e di dipendenza", Biblioteca dell'Economista.

GINI, C. (1926) "The Contributions of Italy to Modern Statistical Methods". Journal of the Royal Statistical Society, Vol. 89, No. 4. (Jul., 1926), pp. 703-724.