Historia del cálculo de probabilidades: de Pascal a Laplace

POR FCO. JAVIER GIRÓN *

1. Los comienzos del cálculo de probabilidades y la estadística

Aunque este artículo versa sobre la historia del cálculo de probabilidades y de la estadística en el siglo XIX, no podemos olvidarnos de sus orígenes, sobre los que vamos a tratar brevemente. ¿Dónde comienza realmente la historia de la estadística? Así comienza un breve artículo de Sir Maurice Kendall (1960), quien concluye diciendo que la estadística en cualquier sentido análogo al nuestro no puede retrotraerse antes de 1660, y señala el trabajo de John Graunt como el punto de partida apropiado. No podemos, ni debemos, separar la historia de la estadística de la del cálculo de probabilidades, ya que, aunque sus fines y sus orígenes sean distintos, no es posible concebir la primera, como ciencia o rama de las matemáticas aplicadas, sin la ayuda de la segunda. La historia así nos lo muestra, sobre todo a partir de mediados del siglo XVIII, donde el pensamiento probabilista se introdujo como una herramienta fundamental para resolver un gran número de problemas aplicados, sobre todo de astronomía y geodesia, sin olvidar los problemas de las hoy llamadas ciencias sociales.

Como es bien sabido, la mayoría de los estudiosos de la historia de la probabilidad atribuyen a los juegos de azar el nacimiento de la teoría de la probabilidad. Dejando a un lado la prehistoria del cálculo de probabilidades, parece que fue Cardano (1526) quien en su *Liber de Ludo Aleae* introdujo ideas probabilísticas en el análisis matemático de los juegos de azar. Posteriormente Tartaglia, Galileo, Pascal, Fermat, Huygens, los Bernoulli, De Moivre y otros contribuyeron al desarrollo de la teoría de la probabilidad a través del análisis de los juegos de azar.

Otros, como por ejemplo Maistrov (1974), opinan que los juegos de azar no fueron el principal estímulo para el origen y desarrollo de la teoría de la probabilidad, sino que dan una explicación materialista de su origen.

---

* Académico Numerario

En los últimos años parece haber resurgido el interés por la historia de la ciencia en general, movimiento al cual la estadística y el cálculo de probabilidades no son ajenos, como lo demuestra el creciente número de libros y monografías publicados en la última década.

Aparte de los ya históricos libros de Todhunter (1865), de David (1962), el ya citado de Maistrov, de los dos volúmenes sobre la historia de la estadística y la probabilidad, editados por Pearson, Kendall y Plackett (1978), del libro de Hacking (1975), del libro editado por Owen (1976) y de los artículos de tipo histórico y biográfico que aparecen a lo largo de los nueve volúmenes de la *Enciclopedia de las Ciencias Estadísticas* editada por Kotz y Johnson a lo largo de los años 1981-1988 y de los de la revista *Biometrika,* podemos citar los más recientes de Stigler (1986), de Porter (1986), los dos volúmenes de Krüger, Daston y Heidelberger (1987), el de Daston (1988), el de Gigerenzer, Swijtink, Daston y Krüger (1989) y el de Hald (1990).

En castellano podemos encontrar breves reseñas históricas en los libros de Cramér (1963), Ríos (1977), Quesada y García Pérez (1985), Peña (1986), DeGroot (1988) y en los artículos de Girón (1989) y, el más reciente, de Fienberg (1992); y es de destacar el trabajo sobre la historia de la estadística en España de Sánchez-Lafuente (1975).

Generalmente se acepta que la historia de la probabilidad comienza en Francia en 1654, año de la correspondencia entre Pascal y Fermat motivada por problemas propuestos por el caballero de Méré, de la que hay una excelente versión anotada en castellano aunque incompleta, en la traducción de las obras completas de Pascal (1981). En ellas por vez primera se abordan el problema de los repartos, el de la duración de una partida y el problema de la ruina en juegos de azar. Su breve *Tratado del triángulo aritmético* tiene interés por la aplicación que de él da al cálculo de combinaciones y su aplicación al problema de determinar los repartos que se deben hacer entre dos jugadores que juegan en varias partidas.

La correspondencia completa entre Pascal y Fermat, de gran valor literario, y la de éste con Carcavi y Huygens se encuentra comentada en la excelente obra sobre los orígenes del cálculo de probabilidades de De Mora (1989).

Como esta correspondencia es bien conocida, aquí me voy a centrar sobre un pasaje de sus *Pensées*: *De la nécésité du pari* o *La apuesta,* aunque su título original es *Infini-Rien* . Su interés es enorme, tanto desde el punto de vista de la Metafísica o de la Religión, como desde el punto de vista de la Teoría de la Probabilidad. De hecho, se trata de una de las primeras tentativas de aplicar el incipiente cálculo de probabilidades a situaciones o problemas extramatemáticos menos frívolos que los planteados por el juego. El arte de conjeturar se transforma aquí en un problema de la *Teoría de la Decisión*, tal como lo interpreta Ian Hacking (1975).

El texto, demasiado largo para incluirlo aquí, que forma parte de los papeles que se encontraron tras su muerte, plantea notables dificultades debido a las correcciones, tachaduras y notas al margen. La formulación del problema planteado por Pascal, básicamente el de si Dios existe o no,

Coeficientes binomiales de Pascal y Fermat. Triángulo de Pascal	1654		
		1662	Tablas de mortalidad de Graunt
Interpretación probabilística de Huygens de la obra de Graunt	1669		
		1693	Estimaciones de Halley de las tablas de vida
Aproximación debida a De Moivre de la distribución binomial por la de Poisson	1712	1713	*Ars Conjectandi* de J. Bernoulli, ley de los grandes números, probabilidad subjetiva y esperanza moral
Teoría de la utilidad de D. Bernoulli	1738	1738	*Doctrina del Azar* de De Moivre. Aproximación normal a la binomial
Libraciones de la luna de Mayer	1750		
		1755	Teoría de errores y sus medias de Sympson
Probabilidad inversa y teorema de Bayes para la binomial	1764		
		1774	Laplace redescubre la probabilidad inversa
Método de la máxima verosimilitud de D. Bernoulli	1778		
		1780	Estimadores de la razón de Laplace
Combinación de ecuaciones de Laplace	1787		
Método de los mínimos cuadrados de Legendre	1805		
		1809	Teoría de errores con distribución normal y mínimos cuadrados de Gauss
Teorema central del límite de Laplace	1810		

Tabla 1. Hitos de la Historia del Cálculo de Probabilidades y de la Estadística. Período (1654-1810).

como un problema de decisión sin experimentación - ya que no se puede aceptar como evidencia ni los milagros ni los testimonios ajenos, que ya supondrían la fe - sería la siguiente: decidir entre dos posibles cursos de acción: vivir de forma piadosa a_1 o vivir de forma mundana a_2. Hay dos posibles estados de la Naturaleza: Dios existe θ_1 o Dios no existe θ_2. Las utilidades correspondientes vendrían dadas por la tabla

	θ_1	θ_2
a_1	u_{11}	u_{12}
a_2	u_{21}	u_{22}

En primer lugar, Pascal establece el siguiente orden en las utilidades: $u_{11} > u_{12} \sim u_{22} > u_{21}$. Un simple argumento, que equivale al del criterio de dominancia, implica que a_1 domina a a_2, lo que, en palabras de Pascal, nos conduciría a apostar por la existencia de Dios. Sin embargo, un libertino nos diría que, si Dios no existe, la vida mundana es preferible a la piadosa, es decir $u_{22} > u_{12}$, con lo cual no hay dominancia entre ambas alternativas. Pascal recurre entonces al criterio de maximizar la esperanza de utilidad. Para ello necesita especificar las probabilidades de los dos estados de la Naturaleza, $p_1 = Pr\{\theta_1\}$, $p_2 = Pr\{\theta_2\} = 1 - p_1$. Pascal, como bien sabemos, no habla en términos de probabilidad sino de apuestas. Comienza con el caso $p_1 = p_2 = 1/2$, y como argumenta que u_{11} es mucho mayor que las otras tres utilidades, entonces la esperanza de utilidad de elegir a_1 es mayor que la de elegir a_2. Como la hipótesis de equiprobabilidad no parece palusible, Pascal invoca, en un pasaje algo confuso, nuevas premisas que equivalen a que, aunque la probabilidad de la existencia de Dios sea muy pequeña (infinitesimal), el orden de infinitud de la utilidad u_{11} es tan grande que siempre $E[a_1; P] > E[a_2; P]$, por lo que, finalmente concluye, la apuesta en favor de la existencia de Dios siempre debe hacerse.

La siguiente contribución importante al cálculo de probabilidades es la publicación en 1657 del primer texto sobre probabilidad *De Ratiociniis in Ludo Aleae*, *Del raciocinio en los juegos de azar*, por el matemático holandés Huygens, motivado por una visita a París en 1655, en la que se interesó por el problema de los puntos. En este libro, por vez primera aparecen alusiones al concepto de esperanza y a las nociones de muestreo con y sin reemplazamiento. Su importancia radicó en popularizar el incipiente cálculo de probabilidades hasta que aparecieron los trabajos más sustanciosos de Pierre de Montmort (1708) *Essai d'analyse sur les jeux de hasard*, la obra póstuma *Ars Conjectandi* de Jacobo Bernoulli (1713) y la *The Doctrine of Chances* de Abraham De Moivre, que someramente analizamos a continuación.

2. La consolidación y extensión del cálculo de probabilidades

En este período, que comprende aproximadamente la primera mitad del siglo XVIII, vemos cómo las incipientes ideas del siglo anterior van formando una teoría más coherente del cálculo de probabilidades.

De la saga de los Bernoulli, los que más contribuyeron al desarrollo del cálculo de probabilidades fueron Nicolás, Daniel y sobre todo Jacobo.

Stigler (1986), con su habitual toque humorístico, en referencia a los Bernoulli, escribe:

> Los Bernoulli son, sin duda alguna, la familia de más renombre en la historia de las ciencias matemáticas. Probablemente hasta doce Bernoulli han contribuído a alguna rama de las matemáticas o la física y, de ellos, al menos cinco han escrito sobre probabilidad. Tan numeroso es el conjunto de los Bernoulli, que simplemente por azar era inevitable que algún Bernoulli debiera considerarse como el padre de la cuantificación de la incertidumbre. El individuo en cuestión es Jacobo Bernoulli (1654-1705), profesor en la Universidad de Basilea desde 1687, contemporáneo y, en ocasiones, rival de Newton.

Antes de analizar la obra de Jacobo Bernoulli, vamos a comentar brevemente la aportación de Nicolás y Daniel Bernoulli a la estadística y al cálculo de probabilidades.

Fue Nicolás quien en 1709 obtuvo la licenciatura en derecho con una disertación en la que aplicaba ideas de la obra de su tío Jacobo *Ars Conjectandi.* Recordemos que fue él quien posteriormente editó esta obra póstuma. Aunque fue un brillante matemático, publicó muy poco. Curiosamente sus logros aparecen ocultos en la correspondencia (unas 500 cartas) con prestigiosos matemáticos de la época como Leibniz, Euler, De Moivre y Montmort. En particular, Nicolás generalizó muchos de los resultados del libro de Montmort. Su amistad fue tan grande que Nicolás pasó largos períodos en casa de Montmort ayudándole a preparar la segunda edición de su libro.

Daniel realizó varias contribuciones a la probabilidad y a la estadística. La primera y la más influyente fue su obra publicada en 1738 *Specimen theoriae novae de mensura sortis* en la que desarrolló el concepto de utilidad.

La idea del artículo se la dió un problema propuesto por su primo Nicolás: el de cómo determinar el precio justo de un juego cuya esperanza matemática es infinita. Generalmente consideramos a la esperanza como el valor justo de un juego, pero como le señaló Nicolás, es injusto pagar una cantidad infinita cuando se está seguro de ganar una cantidad finita. La solución de Daniel fue sustituir el valor monetario del pago por su utilidad al jugador, y calcular su esperanza que bien pudiera resultar finita. Como el artículo de Daniel fue publicado en la revista de la Academia de S. Petersburgo, desde entonces este problema se conoce con el nombre de la paradoja de S. Petersburgo y ha tenido una gran influencia en el desarrollo posterior de la teoría de la utilidad - tras haber estado olvidada durante más de dos siglos - como herramienta para tratar los problemas de decisión en ambiente de riesgo.

El desarrollo posterior de la idea de utilidad comenzó con el trabajo de Frank Ramsay (1926), que desgraciadamente pasó desapercibido en su época hasta que L.J. Savage (1954) lo redescubrió a mediados de nuestro siglo y que hoy, junto con el concepto de probabilidad subjetiva, forma la base de la moderna teoría de la decisión bayesiana. La otra gran contribución a la teoría de la utilidad fue el influyente trabajo de von Neumann y Morgenstern (1944).

Otras contribuciones de Daniel Bernoulli incluyen un contraste de aleatoriedad sobre las órbitas de los planetas y varios artículos, en la década 1770-80, dedicados a desarrollar una teoría probabilística de los errores. Uno de esos artículos es notable, visto retrospectivamente, por utilizar lo que hoy denominamos método de la máxima verosimilitud. Sin embargo, Daniel no desarrolló esta idea y su procedimiento fue duramente criticado como arbitrario por Euler y pronto superado por la teoría bayesiana de Laplace, como veremos más adelante. De modo que la contribución de Daniel Bernoulli no jugó un papel importante en el descubrimiento del método de la máxima verosimilitud por Sir Ronald Fisher en el siglo XX.

La formación filosófica de Jacobo Bernoulli junto con su profunda intuición matemática, le permitieron atacar los problemas conceptuales que aparecen en las aplicaciones del cálculo de probabilidades, tal como lo habían desarrollado sus predecesores, a problemas de probabilidades y evidencia.

Su gran tratado sobre la probabilidad, *Ars Conjectandi* o *El arte de la conjetura* tiene cuatro partes. Las tres primeras están en la línea de sus predecesores sobre todo de Huygens, aunque desarrolla nuevos métodos combinatorios y los aplica a problemas nuevos. Pero es en la parte IV donde desarrolla una nueva teoría de la probabilidad que va más allá de los juegos de azar y que abarca a problemas prácticos de evidencia y decisión. Su diario muestra que había empezado a trabajar en probabilidad hacia 1680 y que había demostrado la ley de los grandes números en 1689. Curiosamente y debido a la tradición filosófica y al pensamiento teológico en que se movió Jacobo Bernoulli, la idea de probabilidad no estaba ligada a la de azar. Ni Pascal, ni Fermat ni Huygens utilizaron la palabra probabilidad en sus escritos. La probabilidad era un atributo de la opinión, un producto de la autoridad. La parte IV de su tratado intentaba salvar las distancias entre los dos conceptos; intentando aplicar la teoría de los juegos de azar a la probabilidad y, por otra parte, preservando la idea de que la probabilidad se basa en argumentos.

Su teoría consta de dos partes: la ley de los grandes números, que podría utilizarse para precisar la fuerza de un argumento y las reglas para combinar evidencias. Estas reglas de combinación eran imperfectas, pero aun hoy siguen siendo interesantes. Algunas de ellas reconocen explícitamente la posibilidad de probabilidades no aditivas. Estas ideas se han retomado de nuevo, notablemente por Shafer, a finales de la década 1970-80, y han servido para construir una teoría de la evidencia y de las funciones de credibilidad con aplicaciones a las nuevas teorías de la inteligencia artificial, sistemas expertos y redes neuronales.

La ley de los grandes números para la distribución binomial aparece al final de su tratado como una solución al problema de determinar una cota inferior al número de observaciones necesarias para estimar una proporción con "certidumbre moral", y lo ilustra con el siguiente ejemplo:

> Para ilustrar ésto con un ejemplo supongo que tengo una urna con 3000 bolas blancas y 2000 bolas negras, que alguien ha puesto allí y que yo, sin conocer tales números, trato de averiguar la proporción de bolas blancas y negras, mediante un experimento reiterado que consiste en sacar una bola al azar, observar su color y reponerla a la urna, con el fin de no disminuir el número de bolas de la urna, y se anota el número total de bolas blancas y negras. La cuestión es, ¿se puede hacer esto de modo que sea diez veces, cien veces, mil veces, etc., más probable (es decir, que sea moralmente cierto) que el número de bolas blancas y negras elegidas estén en la misma razón 3:2 que las bolas de la urna, que estén en otra razón diferente? (Bernoulli, 1713, pp. 225-226).

Bernoulli era consciente del hecho de que la incertidumbre disminuía a medida que aumentaba el número de observaciones y quiso demostrar el principio de que la "certidumbre moral" acerca de una proporción podría aproximarse tanto como se quisiera aumentando el número de observaciones.

El problema planteado por Bernouilli y la solución que propone, lo escribiríamos así, en notación actual: Sea X el número de éxitos observados o casos favorables en N pruebas y sea p la proporción desconocida de éxitos. La solución del problema es que para cualquier número positivo pequeño ε y para cualquier número positivo grande c (por ejemplo, $c = 10$, 100 ó 1000), se puede encontrar un N tal que

$$P\left(\left|\frac{X}{N}-p\right| \leq \varepsilon\right) > cP\left(\left|\frac{X}{N}-p\right| > \varepsilon\right).$$

La versión moderna de lo que hoy llamamos ley débil de los grandes números de Bernoulli se suele escribir de la siguiente manera, que es claramente equivalente a la anterior,

$$P\left(\left|\frac{X}{N}-p\right|>\varepsilon\right)<\frac{1}{(c+1)}\,.$$

Es curioso señalar que en el clásico libro de Uspensky (1937, cap. 6) aparece la demostración original de Bernoulli, que a pesar de ser mucho más larga que la actual basada en la desigualdad de Chebychev, es más natural.

En el problema que Bernoulli se planteó, la proporción p no es cualquier número real, sino el cociente $p = r/(r+s)$, y el ε no era arbitrariamente pequeño sino $1/(r+s)$, para obtener la "certeza moral". El libro de Bernoulli concluye bruscamente con un ejemplo numérico de una cota inferior para N = 25.500, un número muy elevado para el problema concreto que estaba considerando. Se ha especulado con que esto podría ser una de las razones por las que J. Bernoulli no publicó su tratado en vida, más que por sus reservas filosóficas a realizar inferencias basadas en datos empíricos.

El final de este período está marcado, sobre todo, por la obra de De Moivre. La contribución más importante de De Moivre fue la de refinar el teorema de Bernoulli, al obtener la aproximación normal a la distribución binomial, aunque, como comenta Stigler (1986), De Moivre pensó en la curva normal más como un artificio de cálculo que como una distribución continua por derecho propio. Previamente, en 1712, había sugerido la aproximación de la distribución binomial por la de Poisson. Estos resultados, así como el método de las funciones generatrices, se publicaron en la segunda edición de su *The Doctrine of Chances*, en 1718. La deducción de la aproximación normal como una mejora de la cota sobre el tamañomuestral de Bernoulli para obtener la certeza moral pone de manifiesto la importancia de $\sqrt{n}$, donde n es el número de pruebas, como la escala natural respecto de la que se deben medir las desviaciones al centro de la distribución. Sin embargo, De Moivre no fue capaz de ir más allá de la binomial para obtener una generalización del teorema central del límite, ni utilizó su aproximación para hacer inferencias inversas sobre el parámetro p de la binomial.

3. La probabilidad inversa como método de la inferencia estadística

Es curiosa la evolución que el concepto y las aplicaciones del método de la llamada *probabilidad inversa* han sufrido a lo largo del tiempo, desde su origen en 1764 hasta su posición actual como método inferencial en competencia con otros enfoques de la estadística.

Si se relee el capítulo XXXIV del clásico libro de Cramér (1960), sobre regiones de confianza, (no olvidemos que la obra se escribió en los años precedentes a la segunda guerra mundial), el lector, sobre todo si es joven, se verá sorprendido por el autor cuando éste se refiere al antiguo método de Bayes como el "método clásico" para determinar intervalos o regiones de confianza y, en contraposición, aboga por "la moderna teoría" (sic) de regiones de confianza desarrollada por Neymann en los años 30 de este siglo.

Vistos desde nuestra perspectiva, los comentarios de Cramér hoy no nos extrañan ya que, aunque el trabajo de de Finetti sobre probabilidad subjetiva pasó casi desapercibido en los años 20-30 de este siglo, sobre todo en relación con los fundamentos de la inferencia estadística, el resurgimiento de las posturas bayesianas tuvo que esperar hasta la publicación del libro de Savage (1954) y aun así la aceptación de la metodología bayesiana fue lenta, debido posiblemente a la enorme influencia y al prestigio de estadísticos clásicos tan destacados como R.A. Fisher, E.S. Pearson y J. Neymann y de la no menor influencia de los clásicos textos, aparte del de Cramér, de Wilks, Kendall y Stuart y Lehmann, por no citar sino los más influyentes. No es lugar para hablar de la controversia *métodos bayesianos* contra *métodos clásicos*, sino que pasamos inmediatamente a comentar el origen y los personajes asociados al método de la probabilidad inversa.

El primero de ellos es Thomas Bayes cuyo papel en la historia de la probabilidad todavía sigue siendo objeto de controversia. Su famoso artículo *An Essay Towards Solving a Problem in the Doctrine of Chances* ha sido reimpreso cuatro veces en nuestro siglo y sus resultados, redescubiertos como veremos, por Laplace, han sido objeto de debate sobre los fundamentos de la inferencia estadística y el nombre de teorema de Bayes es uno de los epónimos más frecuentes en la literatura estadística moderna.

El reverendo Thomas Bayes murió el 7 de abril de 1761, no el 17 como afirma incorrectamente Stigler (1986) a la edad de 59 años. Se desconoce la fecha de su nacimiento pero muy probablemente fuese en 1701. Aunque fue ministro no conformista en Tunbridge Wells (unos 50 kms. al sureste de Londres, el interés que demostró por las matemáticas le hizo merecedor de ser elegido miembro de la *Royal Society* en 1742. Ninguno de sus trabajos se publicó mientras vivió; sin embargo, uno de sus trabajos póstumos sirvió para que su nombre, dos siglos tras de su muerte, se haya convertido en uno de los epónimos de toda la estadística, a saber, la inferencia bayesiana. Es irónico que el trabajo al que debe su fama, *An Essay Toward Solving a Problem in the Doctrine of Chances*, publicado póstumamente por su amigo y colega Richard Price, tuvo, al parecer, nulo o poco impacto en el desarrollo de la estadística de su tiempo.

El origen del ensayo de Bayes es un misterio. Richard Price, quien encontró el artículo entre los papeles de Bayes, lo leyó en la *Royal Society* el 23 de diciembre de 1763, y fue publicado en las *Philosophical Transactions* de la *Royal Society* en 1764, transcurridos más de dos años y medio de su fallecimiento. Price añadió una sustanciosa introducción y un apéndice al ensayo, de modo que no se sabe exactamente lo que es de Bayes y lo que es de Price. Algunos estudiosos sugieren como explicación del contenido del artículo la de intentar dar una prueba matemática de la existencia de una Primera Causa.

Otra explicación más plausible es la de que la génesis del artículo se debió a la cuestión que Bayes se planteó al leer una carta de Simpson de 1755, que a su vez le indujo a consultar el libro de Simpson de 1740 y la edición de 1738 de la *Doctrine of Chances* de De Moivre (ambos citados en la introducción de Price al ensayo).

En cualquier caso el ensayo está dirigido a resolver el problema de hacer inferencias sobre el parámetro de la binomial, que Jacobo Bernoulli y De Moivre habían dejado pendiente de resolver.

El ensayo es de difícil lectura, debido, en parte, a que adoptó el procedimiento geométrico de Newton, al que sus contemporáneos ingleses estaban más acostumbrados. Pero la dificultad mayor reside en la naturaleza de la probabilidad y en su inversión (probabilidad inversa). Aunque la solución de Bayes es difícil de entender, el *problema* está planteado con toda claridad.

La introducción de Price y el planteamiento sucinto de Bayes, no dejan lugar a dudas sobre el contenido del ensayo:

PROBLEMA

> Dado el número de veces en las que un suceso ha ocurrido y no ha ocurrido: *Se requiere* el azar de que la probabilidad de que ocurra en una sola tirada esté comprendida entre dos grados de probabilidad arbitrarios.

En notación actual el problema lo escribiríamos como sigue: Sea X el número de veces que un suceso ocurre en n pruebas y sea θ la probabilidad de éxito en cada prueba. Se pide calcular $P\,(b < \theta < f \mid x)$.

Bayes procede a su solución de un modo que hoy llamaríamos axiomático, demostrando proposiciones del cálculo de probabilidades a partir de supuestos básicos. Así su Proposición 5 es, básicamente, lo que hoy conocemos como teorema de Bayes. Por medio de un ingenioso mecanismo, conocido como la "mesa de billar" de Bayes, Bayes dedujo, en nuestra notación actual que

$$p\,(b < \theta < f \cap X = p\,) = \int_b^f \binom{n}{p} \theta^p\,(\,1 - \theta\,)^{n-p} d\,\theta\,.$$

Haciendo $b = 0$ y $f = 1$, resulta

$$P\,(\,X = p\,) = \int_0^1 \binom{n}{p} \theta^p\,(1 - \theta\,)^{n-p}\,d\,\theta,$$

integral, que más adelante calcula y cuyo valor es $1/(n+1)$. Por último, utilizando la Proposición 5 obtiene

$$P(b < \theta < f \mid X = p) = \frac{P(b < \theta < f \cap X = p)}{P(X = p)} =$$

$$= \frac{\int_b^f \binom{n}{p} \theta^p (1-\theta)^{n-p} d\theta}{\int_b^f \binom{n}{p} \theta^p (1-\theta)^{n-p} d\theta}.$$

La postura de Bayes, que se refleja en el escolio que sigue inmediatamente a la proposición 9 - su argumento en defensa de una distribución a priori uniforme sobre una probabilidad desconocida - ha sido criticada por Fisher y defendida por Karl Pearson y Harold Jeffreys como un argumento a favor del principio de la razón insuficiente. Sin embargo, recientemente Stigler (1982) ha demostrado que el argumento central del escolio exonera a Bayes del principal defecto del que se le acusaba. Parece claro que Bayes propone el uso de probabilidades subjetivas pero, a ciencia cierta según Stigler, sólo de variables observables y no de variables no observables, como son los parámetros desconocidos.

Del símil de la "mesa de billar", se deduce que la distribución de θ es uniforme. Esto implica que la distribución marginal de la X es también uniforme discreta, ya que

$$P(X = p) = P(0 < \theta < 1 \cap X = p) =$$

$$= \int_0^1 \binom{n}{p} \theta^p (1-\theta)^{n-p} d\theta = \frac{1}{n+1}, \text{ independientemente de } p.$$

Así, para Bayes, el "no saber nada" significaría que en las n pruebas los posibles valores de la observación X serían equiprobables.

Su énfasis en la distribución marginal de observables (la distribución predictiva) lo alinearía con los llamados bayesianos predictivistas tales como de Finetti.

Según Stigler, no fue el problema filosófico de la inversión de la probabilidad lo que le hizo dudar a Bayes de la conveniencia de la publicación de su ensayo, sino el problema de evaluar lo que hoy conocemos como la función beta incompleta, a saber, la integral $\int_0^f \theta^p (1-\theta)^q d\theta$, sobre todo cuando ambos, p y q son grandes.

Fue a Laplace a quien correspondió dar solución al problema de aproximar la distribución beta por la normal, aunque su objetivo no era ese sino probar la consistencia de la distribución a posteriori.

A nuestro segundo personaje, Laplace, se le ha criticado por su adaptabilidad infinita a los distintos cambios de su época. Sus contemporáneos se referían, no sin cierto cinismo, a su "flexibilidad".

Durante los años turbulentos de la revolución francesa mantuvo - literalmente - su cabeza y prosperó. A medida de que sus libros aparecían en sucesivas ediciones, cambiaba las introducciones para ir adecuándose a los tiempos. Así, Laplace dedicó la edición de 1812 de su *Teoría analítica de las probabilidades* a "Napoleón el Grande"; en la edición de 1814 suprimió esta dedicatoria y escribió que "la caída de los imperios que aspiran al dominio universal puede predecirse con probabilidad muy alta por cualquiera versado en el cálculo del azar".

El libro del marqués de Condorcet *Essai sur l'Application de l'Analyse a la Probabilité des Décisions Rendues a la Pluralité des Voix* tiene importancia por tratar de cuestiones sociales y, sobre todo, por servir de estímulo al trabajo fundamental de Pierre Simon de Laplace *Théorie Analytique des Probabilités*, que apareció en sucesivas ediciones en 1812, 1814 y 1820 (esta última con suplementos), y que marca el final de una era. Este libro describe todo el conocimiento hasta la fecha no sólo de la probabilidad sino de la estadística desde una perspectiva muy matemática. Su *Essai Philosophique sur les Probabilités*, que servía como introducción de la obra anterior se reeditó cinco veces entre 1814 y 1825. Sin embargo las contribuciones más importantes de Laplace a la teoría de la probabilidad fueron: el descubrimiento del teorema de Bayes, independientemente de éste, y la utilización de la probabilidad inversa como herramienta de la inferencia estadística; el reconocer la importancia de las transformadas (funciones generatrices y características) y una versión relativamente general del teorema central del límite.

Vamos a comentar, sucintamente, la contribución de Laplace a la solución del problema de la probabilidad inversa.

En su *Mémoire sur la probabilité des causes par les évènements* (1774), leída en la Academia el 31 de mayo de 1780, Laplace comienza con un "Principio", el de la probabilidad inversa y da cuatro ejemplos de aplicación; el principio lo enuncia como sigue:

> Si un suceso se puede producir por un número n de causas diferentes, entonces las probabilidades de esas causas dado el suceso son a cada una como las probabilidades del suceso dadas las causas, y la probabilidad de la existencia de cada una de ellas es igual a la probabilidad del suceso dada esa causa, dividida por la suma de todas las probabilidades del suceso dada cada una de las causas. (Laplace, 1774, 623).

Esto es lo que ahora denominamos teorema de Bayes, cuando las probabilidades de las causas son iguales a priori. Si $A_1, A_2, \ldots, A_n$ son las n causas y E es el suceso, entonces el principio establece que

$$\frac{P(A_i \mid E)}{P(A_j \mid E)} = \frac{P(E \mid A_i)}{P(E \mid A_j)},$$

o

$$P(A_i \mid E) = \frac{P(E \mid A_i)}{\Sigma_j P(E \mid A_j)}.$$

Como el teorema de Bayes afirma

$$P(A_i \mid E) = \frac{P(E \mid A_i)\, P(A_i)}{\Sigma_j P(E \mid A_j)\, P(A_j)},$$

vemos que el principio de Laplace es equivalente al teorema de Bayes cuando las probabilidades a priori son iguales, $P(A_i) = 1/n$.

Aunque el principio parece menos general que el teorema de Bayes, por considerar todas las causas equiprobables, no debemos olvidar que, para Laplace, las causas más probables se podían subdividir en varias causas equiprobables.

Laplace no da una demostración del principio, ya que al parecer lo consideraba como algo natural; sin embargo, en su *Théorie Analytique des Probabilités* (1812, pp. 415-417), aparece una demostración de él.

La primera aplicación de este principio le conduce a demostrar la consistencia de la distribución a posteriori del parámetro de la binomial; es decir, el problema que Bayes no fue capaz de resolver satisfactoriamente, por la dificultad numérica de evaluar la función beta incompleta. Laplace resuelve con éxito este problema utilizando toda la artillería analítica de la que disponía, demostrando, indirectamente, la aproximación normal a la distribución beta, con lo que dejó zanjado el problema originado con Jacobo Bernoulli ochenta años antes.

En esta memoria, Laplace aplica también el procedimiento de la probabilidad inversa al problema de la *elección de medias*, íntimamente relacionado con el problema de la combinación de observaciones, en el que propone como distribución de errores la distribución exponencial doble, también hoy conocida como distribución de Laplace. Una de las dificultades con las que se encontró fue la de que la noción de probabilidad condicionada no estaba suficientemente desarrollada, por lo que sus cálculos no eran del todo correctos.

En una memoria posterior del año 1781, Laplace aplica estos resultados al problema de contrastar la hipótesis de si hay una proporción mayor de nacimientos de varones que de hembras. Utilizando los datos de las memorias de la Academia del año 1771, de que en el período 1745-1770 habían nacido en París 251.527 varones y 241.945 hembras, aplicando sus resultados, obtuvo que la probabilidad a posteriori de que la proporción de varones fuese menor que 1/2 era 1.521×10^{-42}. De aquí, Laplace concluye que es moralmente cierta la hipótesis de una mayor propensión a los nacimientos de varones. Curiosamente, en Londres la preponderancia de varones era aun mayor que en París, donde en el período 1664-1757 se registraron 737.629 niños y 698.958 niñas nacidos. La hipótesis que Laplace formuló fue la de si la probabilidad de nacer varón era la misma en Londres, θ_L, que en París, θ_P. Las frecuencias relativas eran 0,50971 en París y 0,51346 en Londres. A partir de estos datos, Laplace calculó la probabilidad a posteriori

$$P\,(\,\theta_P > \theta_L \mid datos\,) = \frac{1}{410.458}$$

y de este resultado llegó a la conclusión que debía haber alguna causa en Londres que, más que en París, facilitase el nacimiento de los varones; podría depender del clima, de la alimentación o de las costumbres.

A Laplace se le ha criticado frecuentemente por la utilización de distribuciones a priori siempre uniformes, es decir, todas las "causas son igualmente probables". Sin embargo, en la memoria de 1774 analiza un problema sobre lanzamiento de monedas en el que no utiliza una distribución uniforme sobre el todo el espacio paramétrico, sino uniforme en un entorno de 1/2, y en una memoria posterior de 1786 vuelve de nuevo sobre el tema de las distribuciones no uniformes diciendo que aunque se pueden utilizar, de hecho no son necesarias ya que se puede reducir al caso uniforme.

4. El desarrollo del método de los mínimos cuadrados

La historia del desarrollo del método de los mínimos cuadrados está asociada con tres problemas científicos importantes del siglo XVIII, a los que prestaron atención los mejores matemáticos de la época, a saber:

(1) la determinación y representación matemática de los movimientos de la Luna.

(2) la explicación de la desigualdad aparente (no periódica), del movimiento de los planetas Júpiter y Saturno.

(3) la determinación de la forma de la Tierra.

Podemos situar el comienzo de la historia de los mínimos cuadrados alrededor de 1748, con el trabajo de Tobias Mayer sobre las libraciones de la Luna, motivado por el problema de combinar datos u observaciones procedentes de experimentos realizados en diferentes circunstancias. Su problema finalmente consistía en resolver un sistema de 27 ecuaciones con tres incógnitas. Agrupándolas, de modo ingenioso, en tres grupos de nueve ecuaciones y sumándolas en cada grupo obtuvo los estimadores de sus tres parámetros.

Euler en 1748, con motivo de la convocatoria del premio de la Academia de Ciencias de París, sobre el problema de la desigualdad de las órbitas de Júpiter y Saturno, presentó una memoria que resultó ser la ganadora, en la que abordó el problema de combinar ecuaciones de observación. Sin embargo, Euler fracasó en su intento si lo comparamos con el ingenioso método de Mayer.

Posteriormente, Laplace en 1787, en una memoria sobre la desigualdad de los movimientos de Júpiter y Saturno, extendió el método de Mayer, superando los obstáculos con los que se encontró Euler. La mejora de Laplace sobre el método de Mayer fue la de combinar todas las ecuaciones, no solo subconjuntos de ellas, de diversas formas de modo que se aproximó bastante a lo que llamaríamos un método general, que fue popular en la primera mitad del siglo XIX, por ser menos complicado de cálculo que el de los mínimos cuadrados y producir estimaciones bastante precisas.

El nombre del jesuita Boscovich aparece en escena, en 1755, en relación con el problema de determinar la forma de la Tierra, enunciando una solución al problema en forma de un *principio*, es decir una lista de propiedades que debería poseer la *media* basada en una combinación de observaciones; el principio se publicó posteriormente en 1760.

La importancia del trabajo de Boscovich fue sobre todo, aparte de su interés por formular el problema en términos de un principio, que hizo que Laplace se volviera a interesar en el problema de la determinación de la forma de la Tierra. En 1789 propuso el llamado "Método de Situación" que supuso un pequeño adelanto respecto del método de Boscovich. Diez años después, mientras preparaba el segundo volumen de su *Mécanique céleste,* modificó de nuevo el principio de Boscovich. La principal conclusión de este trabajo, que supuso un avance significativo en la técnica estadística, fue la de introducir unas ponderaciones en las ecuaciones de acuerdo con la precisión de las observaciones.

Pero la solución al problema tuvo que esperar hasta el comienzo del siglo XIX, y vino de la mano de Legendre.

Legendre (1752-1833) fue un matemático de gran originalidad y de, lo que hoy llamaríamos, amplio espectro. Tres años más joven que Laplace, le sucedió en sus cargos de catedrático de matemáticas en la Ecole Militaire y en la Ecole Normale. Sus trabajos más conocidos son los dedicados a las integrales elípticas, a la teoría de números y a la geometría. Su libro *Eléments de géometrie* fue uno de los textos más famosos, en esta especialidad, del siglo XIX. Fue miembro de dos comisiones francesas, una que en 1787 unió geodésicamente los observatorios de París y Greenwich y otra que en 1795 midió el arco del meridiano de Barcelona a Dunquerque, arco sobre el que se basaba la definición de longitud del metro. Precisamente esto resultó ser el nexo de unión entre los trabajos sobre astronomía y geodesia y el método de los mínimos cuadrados.

En 1805 Legendre publicó el trabajo *Nouvelles méthodes pour la déterminacion des orbites des comètes*. Curiosamente resutó ser el apéndice al trabajo el que le confirió un lugar importante en la historia de la estadística. El objetivo que se propuso lo definió claramente como sigue:

Sobre el método de los mínimos cuadrados

> Muchas investigaciones donde el objetivo es deducir los resultados más precisos de las medidas observacionales, nos conducen a un sistema de ecuaciones de la forma
>
> $$E = a + bx + cy + fz + \&c.,$$
>
> en las que a, b, c, &c, son coeficientes conocidos, que varían de una ecuación a otra, y x, y, z, &c., son cantidades desconocidas, que hay que determinar con la condición de que cada valor de E se reduzca a cero o a una cantidad muy pequeña. (Legendre, 1805, p. 72).

A continuación, cuando el número de ecuaciones es mayor que el de incógnitas, Legendre propone como solución la siguiente:

> De todos los principios que se pueden proponer, pienso que no hay otro más general, más exacto, o más fácil de aplicar que el qLe hemos utilizado en este trabajo; consiste en hacer *mínima* la suma de los cuadrados de los errores. Con este método se establece una forma de equilibrio entre los errores que, ya que impide que los errores extremos dominen, es apropiado para revelar el estado del sistema que mejor se aproxima a la realidad. (Legendre, 1805, pp. 72-73).

Inmediatamente, calculando el mínimo por derivación, deduce las que hoy llamamos "ecuaciones normales" (nombre acuñado más tarde por Gauss),

$$0 = \int ab + x \int b^2 + y \int bc + z \int bf + \&c.$$

$$0 = \int ac + x \int bc + y \int c^2 + z \int fc + \&c.$$

$$0 = \int af + x \int bf + y \int cf + z \int f^2 + \&c.$$

que, en palabras de Legendre, se resuelven por "métodos ordinarios".

Las ventajas del método eran evidentes; a las que Legendre añadió las siguientes:

(1) Si había un ajuste perfecto, su método lo encontraba.

(2) Si posteriormente había que desechar una ecuación, simplemente se revisaban las ecuaciones normales, restándolas los términos correspondientes.

(3) La media aritmética era un caso particular del método, cuando sólo hay una incógnita con coeficientes constantes $b = b' = \ldots = 1$.

(4) Calcular el centro de gravedad de varias masas no era sino un caso especial del método.

La claridad de su exposición contribuyó, sin duda, al hecho de que el método fuese aceptado con entusiasmo tras su publicación. Diez años más tarde el método era una herramienta estándar en astronomía y geodesia en Francia, Italia y Prusia y poco después llegó a Inglaterra. El método tuvo tal difusión que pronto relegó al olvido a los demás métodos.

Con la introducción de los mínimos cuadrados por Legendre se llegó al final de una primera etapa partiendo de las líneas comenzadas, por separado, por Mayer y Boscovich medio siglo antes. La idea de combinar diferentes ecuaciones de observación se originó con los trabajos astronómicos de Mayer, mientras que la idea de buscar un criterio objetivo de ajuste partió del trabajo geodésico de Boscovich y fue Legendre quien sintetizó ambas corrientes. Sin embargo algo importante faltaba. Ni el procedimiento tenía justificación probabilística ni había posibilidad de cuantificar el error de los estimadores. Resulta curioso como Laplace, que había jugado un papel importante en el desarrollo inicial del método, tuvo que esperar hasta que se publicó el trabajo de Gauss de 1809, para llegar finalmente a lo que Stigler (1986) denomina la síntesis de Gauss-Laplace.

En 1809 Gauss publica su *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium.* Al final del libro añadió una sección dedicada al problema de la combinación de observaciones. Pero había una profunda diferencia en su enfoque con respecto al de sus predecesores. Gauss resolvió el problema en términos propiamente probabilísticos. Por primera vez supuso que los posibles valores de los errores eran variables aleatorias con una distribución $\varnothing(\Delta)$ y, para resolver el problema de combinar ecuaciones u observaciones cuando hay más ecuaciones que incógnitas, utilizó la versión de Laplace del teorema de Bayes para calcular la distribución a posteriori, usando una distribución uniforme para los parámetros desconocidos para, posteriormente, maximizar la densidad a posteriori, anulando las derivadas y resolviendo el sistema de ecuaciones resultante. Pero las ecuaciones resultantes dependían de la forma de la distribución común de los errores $\varphi(\Delta)$.

Gauss, en vez de imponer condiciones sobre la distribución $\varphi(\Delta)$ que la caracterizasen, tales como que, por ejemplo, tuviese un máximo en $\Delta = 0$, fuese simétrica, se anulase fuera de un posible rango de los posibles errores, etc., procedió al revés. Supuso que

$$\varphi(\Delta) = \frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2}$$

ya que esta distribución, la normal como se la conoció posteriormente, conducía a la media aritmética de las observaciones en el caso más simple.

Gauss demostró a continuación, que, en el caso general, esta distribución conducía efectivamente al método de los mínimos cuadrados.

En abril de 1810 Laplace presentó en la Academia su teorema central del límite y, al parecer como demostraron los acontecimientos posteriores, no había llegado todavía a sus manos el libro de Gauss, ni había pensado en la curva normal como una distribución posible para los errores. Ya habíamos comentado, al hablar de su contribución a la probabilidad inversa, su propuesta de la distribución exponencial doble como distribución de los errores.

Al final de ese año, inmediatamente tras haber leído el libro de Gauss, a Laplace se le hizo evidente la conexión entre su teorema central del límite y la estimación lineal producida por el método de los mínimos cuadrados. En efecto, si los errores de la formulación de Gauss eran agregados de un número elevado de pequeños errores, el teorema central del límite implicaría que estarían distribuidos aproximadamente como la curva de Gauss, $\varphi(\Delta)$, (este argumento de Laplace se conocería más tarde con el nombre de la hipótesis de los errores elementales).

Como ocurrió con otras célebres controversias de la historia de la matemática, la más conocida fue sin duda la disputa entre Newton y Leibniz sobre el origen del cálculo infinitesimal, el descubrimiento del método de los mínimos cuadrados produjo el enfrentamiento entre Gauss y Legendre. Gauss se refirió al método de los mínimos cuadrados como "nuestro principio" argumentando que venía utilizándolo desde el año 1795. Parece ser, según recientes investigaciones de Stigler, que aunque Gauss tuviese el método antes que Legendre, no se dió cuenta de su potencialidad hasta que apareció la publicación de Legendre.

5. Epílogo

Laplace murió el 5 de marzo de 1827, poco antes de cumplir 75 años, y su muerte marcó el fin de una era y el principio de la llamada *Revolución probabilística*, que será el contenido de la conferencia del Prof. Ríos, dentro de este segundo ciclo de la *Historia de la Matemática del siglo XIX.*

Las bases matemáticas del moderno cálculo de probabilidades quedaban firmemente establecidas en la *Théorie Analytique des Probabilités* de Laplace y por esas fechas la aplicación del método de los mínimos cuadrados - tras la síntesis de los trabajos de Gauss y Laplace sobre la combinación de observaciones y la distribución normal (que en la literatura francesa se continúa llamando ley de Gauss-Laplace) -, era común en ciertas ciencias aplicadas como la geodesia y la astronomía pero permaneció casi por completo desconocido en las ciencias sociales. Hubo que esperar bastante tiempo hasta que las nuevas ideas estadísticas se abrieron camino en un número cada vez mayor de aplicaciones.

6. Agradecimientos

Quiero expresar mi agradecimiento al Prof. Miguel A. Gómez Villegas por su desinteresada ayuda en la preparación de este artículo, y por proporcionarme un buen número de fuentes bibliográficas.

Este trabajo se ha realizado con ayuda parcial de la *Consejería de Educación de la Junta de Andalucía* como parte del *Plan Andaluz para la Consolidación de Grupos de Investigación y Desarrollo Tecnológico.*

Referencias

BAYES, T. (1763). An Essay towards Solving a Problem in the Doctrine of Chances, with Richard Price's Foreword and Discussion. *Phil. Trans. Roy. Soc.*, pp. 370-418.

BERNOULLI, D. (1738). Specimen theoriae novae de mensura sortis. *Comentarii Academiae Scientiae Imp. Petropolitanae.* V, 175-192.

BERNOULLI, J. (1713). *Ars Conjectandi.* Impensis Thurnisiorum: Fratrun, Basilea.

CARDANO, G. (1526). *Liber de Ludo Aleae.*

CONDORCET, M.J.A.C. (1785). *Essai sur l'Application de l'Analyse a la Probabilité des Décisions Rendues a la Pluralité des Voix.* París.

CRAMÉR, H. (1960). *Métodos Matemáticos de Estadística.* Aguilar: Madrid.

CRAMÉR, H. (1963). *Elementos de la Teoría de Probabilidades y aplicaciones.* Aguilar: Madrid.

DASTON, L.J. (1988). *Classical Probability in the Enlightenment.* Princeton Univ. Press.

DAVID, F.N. (1962). *Games, God and Gambling.* Charles Griffin: London.

DEGROOT, M.H. (1988). *Probabilidad y Estadística.* Addison-Wesley Iberoamericana: México.

FIENBERG, S.E. (1992). The history of statistics: a review essay. *Statistical Science,* **7**, 208-225.

GIGERENZER, G., SWIJTINK, Z., PORTER, T., DASTON, L., BEATTY, J. and KRÜGER, L. (1989). *The Empire of Chance. How Probability Changed Science and Everyday Life.* Cambridge Univ. Press.

GIRÓN, F.J. (1989). El concepto de probabilidad. En *Historia de la Ciencia Estadística.* Real Acad. Cien. Exac., Fís. y Nat. Madrid.

HACKING, I. (1975). *The Emergence of Probability.* Cambridge University Press: Cambridge.

HUYGENS, C.I. (1920). Ouvres Completes, Vol. 14. Martinus Nijhoff: La Haye.

HALD, A. (1990). *A History of Probability and Statistics and Their Applications before 1750.* Wiley, New York.

KOTZ, S. and JOHNSON, N.L. (1982). *Encyclopedia of Statistical Sciences. Volume 1.* John Wiley: New York.

KOTZ, S. and JOHNSON, N.L. (1982). *Encyclopedia of Statistical Sciences. Volume 2*. John Wiley: New York.

KOTZ, S. and JOHNSON, N.L. (1983). *Encyclopedia of Statistical Sciences. Volume 3*. John Wiley: New York.

KOTZ, S. and JOHNSON, N.L. (1983). *Encyclopedia of Statistical Sciences. Volume 4*. John Wiley: New York.

KOTZ, S. and JOHNSON, N.L. (1985). *Encyclopedia of Statistical Sciences. Volume 5*. John Wiley: New York.

KOTZ, S. and JOHNSON, N.L. (1985). *Encyclopedia of Statistical Sciences. Volume 6*. John Wiley: New York.

KOTZ, S. and JOHNSON, N.L. (1986). *Encyclopedia of Statistical Sciences. Volume 7*. John Wiley: New York.

KOTZ, S. and JOHNSON, N.L. (1988). *Encyclopedia of Statistical Sciences. Volume*. 8. John Wiley: New York.

KOTZ, S. and JOHNSON, N.L. (1988). *Encyclopedia of Statistical Sciences. Volume* 9. John Wiley: New York.

KRÜGER, L., DASTON, L. and HEIDELBERGER, M., eds. (1987). *The Probabilistic Revolution. Volume 1: Ideas in History*. MIT Press.

KRÜGER, L., GIGERENZER, G. and MORGAN, M.S. eds. (1987). *The Probabilistic Revolution. Volume 2: Ideas in the Science*. MIT Press.

LAPLACE, P.S. (1812). *Thèorie Analytique des Probabilités*. Courcier: París.

MAISTROV, L.E. (1974). *Probability Theory: A historical Sketch*. Academic Press: New York.

DE MOIVRE, A. (1740). *The Doctrine of Chances*. London. (Reprinted in 1967, by Chelsea, New York).

DE MORA CHARLES, M. (1989). *Los Inicios de la Teoría de la Probabilidad: SiglosXVI y XVII*. Servicio Editorial. Univ. del País Vasco. Vizcaya.

MONTMORT, P. (1708). *Essai d'Analyse sur les Jeux de Hasard*. Quilau: París.

PASCAL, B. (1981). *Obras: Pensamientos. Provinciales. Escritos científicos. Opúsculos y cartas*. Ediciones Alfaguara: Madrid.

PEÑA, D. (1986). *Estadística. Modelos y Métodos: 1. Fundamentos*. Alianza Editorial: Madrid.

PEARSON, E.S. and KENDALL, M. (1978). *Studies in the History of Statistics and Probability. Volume 1* (2nd impression). Charles Griffin: London.

PORTER, T.M. (1986). *The Rise in Statistical Thinking, 1820-1900*. Princeton Univ. Press.

QUESADA, V. y GARCÍA PÉREZ, A. (1985). *Curso Básico de Probabilidades*. ICE: Madrid.

RAMSAY, F.P. (1926). Truth and probability. Reprinted in *Studies in Subjective Probability, 2nd edition, 1980*. (H.Kyburg and Smokler, eds.), pp. 23-52. Wiley: New York.

RÍOS, S. (1977). *Métodos Estadísticos*. Castillo: Madrid.

SÁNCHEZ-LAFUENTE, J. (1975). *Historia de la Estadística como Ciencia (1500-1900)*. Instituto Nacional de Estadística: Madrid.

SAVAGE. L.J. (1972). *The Foundations of Statistics.* Dover: New York.

STIGLER, S.M. (1982). Thomas Bayes's Bayesian inference. *J. Roy. Statist. Soc. A.* **145,** 250-258.

STIGLER, S.M. (1986). *The History of Statistics: The Measurement or Uncertainty before 1900.* Harvard Univ. Press.

SHAFER, G. (1976). *A Mathematical Theory of Evidence.* Princeton University Press: Princeton, N.J.

TODHUNTER, I. (1985). *History of the Mathematical Theory of Probability.* Cambridge University Press: London. Reprinted by Chelsea, New York, 1949, 1961.