Historia de la Probabilidad y la Estadística [V],(J.M. Riobóo e I. Riobóo, eds.) Santiago de Compostela: Nino-Centro de Impresión Digital
(2011), 225-235.

CÓMO LA APROXIMACIÓN BAYESIANA A LA INFERENCIA PASÓ A SER FRECUENTISTA

Miguel A. Gómez Villegas
Departamento de Estadística e Investigación Operativa
Fac. de Ciencias Matemáticas
Universidad Complutense
Plaza de Ciencias, 3
28040-MADRID
Tfno. 913944428 Email: ma.gv@mat.ucm.es

En este trabajo se va a poner de manifiesto cómo la primera aproximación que se hizo a un problema de inferencia fue bayesiana, y posteriormente se verá el motivo y los personajes que hicieron que la aproximación pasara a ser frecuentista.

La inferencia estadística se puede decir que surge con la figura de Thomas Bayes (1701?-1761), tres años después de su muerte, en 1764, ve la luz su trabajo titulado *An Essay towards solving a problem in the Doctrine of Chances*, publicado por Richard Price. Un estudio pormenorizado de este Ensayo puede verse en Gómez Villegas (2001). La primera traducción del *Ensayo* al castellano, puede verse en Girón y otros (2001), que también contiene copia de un facsímil del original que se conserva en la Biblioteca de la Real Academia de Ciencias de Madrid.

El antecedente más directo que se puede citar del trabajo de Bayes abordado en el Ensayo, es el trabajo de James Bernoulli (1654-1705) titulado *Ars Conjectandi*, que contiene la distribución de Bernoulli y en el que se introduce el concepto de la *esperanza moral* de un suceso, como un intento de determinar la probabilidad de un suceso asociado a un experimento del tipo éxito o fracaso, en función del número de veces que ha sido observado el suceso en n repeticiones del experimento. Este problema es recogido por De Moivre (1667-1754). Este hugonote francés que se ve obligado a exilarse a Inglaterra, trata el problema en su libro *Doctrine of Chances.* Con notación actual, se puede decir que intuye que el cociente r/n, donde r es el número de veces que se ha observado el suceso en n repeticiones, se debe aproximar a la probabilidad de éxito θ en la distribución de Bernoulli.

Pero tanto Bernoulli como De Moivre, hacen aseveraciones acerca de lo que en la época se llamaban *problemas directos* de probabilidad; problemas en los que se supone conocida la probabilidad de éxito θ y se calcula la probabilidad de cualquier sucesión de éxitos y fracasos.

Por el contrario, el problema que abordó Bayes en su Ensayo es el que se llamó desde entonces *problema inverso.* Se trata realmente del problema de inferencia, de conocer θ la causa, observado su efecto r. En expresión actual, lo que hace Bayes es calcular

$$P\{a < \theta < b | X = r\}$$

siendo X una variable con distribución binomial de parámetros n y θ. Para dar sentido a esta expresión, Bayes necesita hacer dos cosas; introducir una probabilidad inicial sobre θ y definir la densidad final $\pi(\theta|r)$. Es notable que Bayes diera estos dos pasos y pone de manifiesto la alta capacidad matemática de Bayes. El descubrimiento fue hecho otra vez unos años más tarde, y con independencia, por Laplace (1794-1827) quién en su *Theorie Analitique des Probabilities*, obtiene además la distribución final para modelos más generales que el binomial.

Para resolver el problema de introducir una probabilidad inicial sobre θ Bayes interpreta la probabilidad como:

> *La probabilidad de cualquier suceso es el cociente entre el valor en el que uno espera dependiendo de la ocurrencia del suceso que debe ser calculado, y el valor de la cosa esperada una vez que ésta ha ocurrido.*

La definición no es muy clara, pero lo que si que es cierto es que se trata de una aproximación subjetiva a la probabilidad.

Lo que hoy se conoce como versión continua de la fórmula de Bayes, está contenido en la proposición 9 de la Sección II del Ensayo, y no es otra cosa que la expresión de la probabilidad condicionada como cociente entre la densidad conjunta y la marginal; en notación actual

$$P\{a < \theta < b | X = r\} = \frac{\int_a^b \binom{n}{r} \theta^r (1-\theta)^{n-r} d\theta}{\int_0^1 \binom{n}{r} \theta^r (1-\theta)^{n-r} d\theta}$$

Tomando cómo distribución inicial para θ la uniforme en el intervalo $[0,1]$ y el modelo binomial, mediante las expresiones

$$\pi(\theta) = I_{[0,1]}(\theta) \qquad P\{X = r|\theta\} = \binom{n}{r} \theta^r (1-\theta)^{n-r},$$

Bayes calcula la densidad final por la fórmula

$$\begin{aligned} \pi(\theta | X = r) &= \frac{\pi(\theta) P\{X = r|\theta\}}{\int_0^1 \pi(\theta) P\{X = r|\theta\} d\theta} \\ &= \frac{\theta^r (1-\theta)^{n-r}}{\int_0^1 \theta^r (1-\theta)^{n-r}}. \end{aligned}$$

El paso que Bayes da es de gigante, capta con precisión el problema de estimación sobre la probabilidad de éxito para la distribución binomial, calcula la expresión de la distribución final, e introduce la distribución uniforme como distribución inicial. Comentarios sobre estos aspectos pueden verse en Gómez Villegas (2001).

El siguiente estadístico en la historia es tambien bayesiano, se trata de Laplace. En 1774 lee ante la Academia francesa su *Mémoire sur la probabilité des causes par les évènements* y en ella enuncia un principio, que es la versión discreta del teorema de Bayes:

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

Esto es lo que ahora denominamos teorema de Bayes o versión discreta del teorema de Bayes, cuando todas las causas son inicialmente equiprobables.

Si $C_1, C_2, ..., C_n$ son las n causas y E es el suceso, las causas son sucesos disjuntos y su unión da el total, se tiene que

$$\frac{P(C_i|E)}{P(C_j|E)} = \frac{P(E|C_i)}{P(E|C_j)} \qquad o \qquad P(C_i|E) = \frac{P(E|C_i)}{\sum_{j=1}^{n} P(E|C_j)}$$

Como el teorema de Bayes dice que

$$P(C_i|E) = \frac{P(C_i)P(E|C_i)}{\sum_{j=1}^{n} P(C_j)P(E|C_j)},$$

cuando las probabilidades de las causas son todas iguales, es decir, $P(C_i) = \frac{1}{n}$, las probabilidades se simplifican. Aunque el teorema parece menos general que el de Bayes, hay que tener en cuenta que, para Laplace, las causas siempre se pueden subdividir en varias causas equiprobables. La demostración que Laplace hace de este resultado está contenida en sú *Teoría Analítica*, página 182, y es exactamente igual que la que figura en nuestros libros de texto. También parece fuera de toda duda que el propio Bayes conocía este teorema, su versión continua es una generalización, y para su obtención forzosamente tuvo que basarse en él.

También incluye Laplace en su *Teoría Analítica* (página 364), la versión continua de la fórmula de Bayes, cuando la distribución inicial no es constante, mediante la expresión

$$\pi(\theta|\mathbf{x}) = \frac{\pi(\theta)f(\mathbf{x}|\theta)}{\int \pi(\theta)f(\mathbf{x}|\theta)d\theta}$$

Despues de Laplace, la probabilidad se aplica a la *aritmética social*, lo que llamaríamos actualmente sociología, con las figuras de Condorcet, Quetelet y Arbuthnot, y es la probabilidad inversa, o la aproximación bayesiana, el método indiscutible de estimación, hasta que Boole publica su libro *Laws of Thought* (1854), en el que introduce su método simbólico de inferencia lógica y pone por primera vez en entredicho la aproximación bayesiana. Básicamente las críticas provienen de considerar a los valores desconocidos como parámetros y no cómo variables aleatorias. En ese caso ya no es necesario asignar distribuciones de probabilidad iniciales dependientes de cada situación y de cada individuo.

Las figuras de Karl Pearson, Fisher, Neyman, y Egon Pearson

Karl Pearson (1857-1936) nace en Londres, estudia en el *University College School* y matemáticas en el *King College* de Cambridge. Con 22 años marcha a Alemania a estudiar leyes, fśica y metafísica. De 1880 a 1884 es profesor de matemáticas en el *King College* y en el *University College*. En 1911, a la retirada de Galton, ocupa la cátedra Galton de eugenesia. Pearson era un darwinista convencido y un ferviente socialista. Entre 1892 y 1892 da conferencias en el *Greshan College* sobre geometría de la estadística. Estas conferencias marcan el comienzo de una época en la teoría y la práctica de la estadística, a partir de este momento se puede admitir que la estadística ha alcanzado su edad adulta.

Entre 1893 y 1906 publica 100 artículos sobre aplicaciones de la estadística, este trabajo se ve acompañado por la fundación de la revista *Biometrika* para recoger publicaciones de estadística aplicada a la Biología. En 1905 publica *On the general theory of skew correlation and non linear regression* donde introduce el concepto de coeficiente de correlación lineal y generaliza la idea de la independencia o dependencia en estadística. En 1925 funda la revista *Annals of Eugenics*, en 1932 se retira del *University College*, que divide la cátedra de estadística en dos: una de eugenesia, que desempeñará Fisher y una de estadística, que desempeñará Egon Pearson. Todavía en 1934 publica *The Tables of the Incomplete Beta-Function* con las que resuelve el problema de la aproximación de la probabilidad final que había encontrado Bayes en su *Ensayo.*

Para hacerse una idea de su carácter, 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 envíe sus fondos a la *Academie des Sciences* para la fundación de un laboratorio de probabilidad que aplique ésta al problema de la evolución biológica.

Entre sus contribuciones más notables cabe citar la introducción de la familia de curvas que contiene a la beta, la gamma y la normal, para modelizar fenómenos aleatorios, el método de los momentos para ajustar los parámetros de las distribuciones, el desarrollo de la correlación y de la regresión y la introducción del método de la χ^2 de Pearson, en 1900, para dar una medida del ajuste entre un modelo y una muestra.

El otro estadístico que contribuye al desarrollo de los métodos frecuentistas es Ronald Aylmar Fisher (1890-1962). Fisher nace en East Finley (Londres)y estudia en Stanmore y Harrow; en su juventud tuvo prohibido el estudiar con luz artificial, debido a su delicada visión. Gracias a una beca estudió en el *Casius College* de Cambridge, dónde se graduó entre 1909 y 1912.

En 1913 fue lector de física matemática, cursó estudios de biometría y genética, para descubrir de 1913 a 1915 que su vocación no era trabajar en una compañía de inversiones. En 1916 escribe un artículo demostrando que las teorías de Mendel no son rechazadas por la evidencia estadística. El artículo lo referencia K. Pearson como estadístico y Punnet cómo genético y al no lograr publicarlo en *Biometrika*, la revista dirigida por K. Pearson, constituyó el motivo de su enemistad con él. En 1917 se casa con Ruth E. Guinnes con quien tuvo dos hijos y seis hijas.

En 1919 se une a la estación experimental de Rothamsted donde desarrolla el análisis de la varianza y el diseño de experimentos. En 1924 es elegido miembro de la Royal Society y en 1930 publica *The Genetical Theory of Natural Selection*, donde apoya y actualiza la teoría de la evolución de Darwin. En 1925 publica *Statistical Methods for Research Workers*, donde recoge una gran cantidad de los conceptos sobre los que se basa la aproximación frecuentista a la inferencia, bastantes de los cuales son debidos a Fisher. En 1933 acepta la cátedra de eugenesia en el University College de Londres y se dedica a trabajar a fondo en genética.

Este trabajo le lleva a publicar en 1935 *The Design of Experiments*, que contiene las ideas de Fisher respecto a cómo analizar experimentos desde el punto de vista estadístico, una crítica de la probabilidad inversa, el famoso experimento del *tea-testing* y aspectos interesantes en defensa de los diseños factoriales. Entre 1938 y 1954 se dedica a inferencia estadística. Entre 1952 y 1954 es presidente de la Royal Statistical Society, dedicando sus intervenciones a glosar las contribuciones de los primeros estadísticos. En 1956, con 66 años, publica *Statistical Methods and Scientific Inference*, un libro en el que están recogidas sus principales contribuciones estadíticas.

En 1957 Fisher se retira y se marcha a Adelaida, Australia, como investigador senior. Allí muere de cáncer de boca a la edad de 72 años.

Es interesante recoger los comentarios, expuestos en el libro *The Design of Experiments* (1990 pág. 6), que le llevan a Fisher a estar en contra de la aproximación bayesiana.

El primer argumento es que el teorema de Bayes no fué publicado por éste; si hubiera sido una cosa a la que el autor le hubiera dada gran importancia debería de haber sido publicada por él mismo. El segundo argumento es que el teorema de Bayes no supone la probabilidad cómo una cantidad objetiva, sino cómo algo que mide tendencias psicológicas. Y añade: un axioma que sea cierto debe ser admitido automáticamente por todo aquel a quién se le muestre; lo cuál no ocurre con la aproximación bayesiana. (Aunque personalmente no comparto los argumentos de Fisher en contra de la aproximación bayesiana, considero importante el recogerlos).

Un estudio detallado de las contribuciones científicas de Fisher a la estadística puede verse en Girón y Gómez Villegas (1998).

El desarrollo de los métodos estadísticos bayesianos y frecuentistas, puede verse en el libro de Gómez Villegas (2005).

No se van a dar muchos datos de las biografías de Neyman y Egon Pearson, para no alargar esta contribución, Neyman nace en Bendery, en Rusia entonces y ahora en Moldavia. La contribución de la familia de Neyman a la causa polaca, le valió a Neyman ser varias veces encarcelado; en un intercambio de prisioneros es trasladado a Varsovia, donde conoce a Sierpinsky quién le busca trabajo para asistir en experimentos agrícolas con sus conocimientos de estadística. Este trabajo le hace solicitar una beca al *University College* de Londres para trabajar con Karl Pearson; allí conoce a Egon Pearson con quién desarrollará la teoría de los tests de hipótesis y de los intervalos de confianza. Estas investigaciones van a llevar definitivamente a la inferencia estadística por la senda frecuentista, alejándola de la aproximación bayesiana inicial.

El renacimiento bayesiano

El resurgir de las ideas bayesianas se debe principalmente a las figuras de Jeffreys (1891-1989) y de Finetti (1906-1985).

Harold Jeffreys es de formación básica un astrónomo no un estadístico. Nació en Fatfield, en el condado de Durham en Inglaterra. Estudió en el *St John College* en Cambridge, posteriormente explicó matemáticas, geofísica y llegó a ser en Cambridge, catedrático de astronomía. Se casó con una matemática y física, Bertha Swirles, en 1940.

Jeffreys desarrolló una teoría lógico deductiva de la probabilidad, para contrastar hipótesis científicas y para resolver problemas que se presentan, relacionados con el análisis de datos, en las ciencias experimentales. Las ideas de Jeffreys están contenidas en su libro *Theory of Probability* y consisten en admitir ciertos axiomas que llevan al cálculo de probabilidades y al teorema de Bayes. Rechaza la interpretación frecuentista de la probabilidad, deduce lo que se conoce como *regla de Jeffreys* para representar poca información inicial, basándose en consideraciones de invariancia:

$$\pi(\theta) = \sqrt{I_n(\theta)},$$

donde $I_n(\theta)$ es la llamada *información de Fisher* para muestras aleatorio simples de tamaño n: la esperanza del cuadrado de la derivada respecto al parámetro del logaritmo de la función de verosimilitud. Esta distribución inicial, tiene la propiedad de que permanece igual, aunque se haga una transformación de la parametrización (técnicamente una transformación regular).

Introduce Jeffreys también el concepto de *factor Bayes* para los contrastes de hipótesis, e introduce distribuciones mixtas para contrastar hipótesis que tengan menor dimensión que la dimensión del espacio paramétrico.

El otro estadístico, que contribuye a que la aproximación bayesiana vuelva a ocupar el valor que le corresponde, es Bruno de Finetti. Nace en Innsbruck (Austria) en 1906 de padres italianos; su padre era ingeniero y trabajaba en la construccion del ferrocarril. Estudió primaria y secundaria en Trento, lugar de origen de su madre, a donde se trasladó con ella, tras la prematura muerte de su padre. Ingresó posteriormente en el recientemente creado Instituto Politécnico de Milán, pero tras dos años de estudios, se transladó a la Universidad de Milán, por la que se licenció en matemáticas en 1927. En una carta a su madre escribe :

> *Las matemáticas no son un campo totalmente explorado, sólo para ser aprendidas y pasar a la posteridad como están, siempre están en progreso, enriqueciéndose e iluminándose, son una criatura viva y vital, en total desarrollo y por estas razones yo las amo, las estudio y deseo dedicar mi vida a ellas.*

Un primer trabajo de de Finetti sobre herencia mendeliana (en 1926) llamó la atención de Conrado Gini, quien le propuso como director de la oficina matemática del *Instituto Centrale di Statistica* en Roma. Los cuatro primeros años en la capital italiana, de 1927 a 1931, constituyen la época en la que aparecen sus grandes aportaciones originales, contemporáneas de las de Kolmogorov, Levi, Fisher o Cantelli. Así en 1928 presenta *Funzione Caratteristica di un Fenomeno Aleatorio* ante el Congreso Internacional de Matemáticas de Bolonia, un trabajo que contiene su concepto de intercambiabilidad; en la audiencia destacan Borel, Cantelli, Darmois, Fisher, Fréchet, Khintchine, Lévi, Neyman y Polya. En 1946 obtiene una cátedra en la Universidad de Trieste. Su estancia en Trieste, de 1931 a 1954, coincide con el desarrollo de su imagen internacional. Su trabajo más conocido es, *La Prevision, ses Lois Logiques, ses Sources Subjectives* (1937), que contiene el desarrollo matemático formal de la probabilidad como grado de creencia y el teorema de representacion de variables dicotómicas intercambiables. En 1950 es invitado al segundo Simposio de Berkeley, donde presenta un trabajo sobre métodos bayesianos robustos. En 1952, es invitado a visitar Chicago; es el principio de una larga y fructífera colaboración con Leonard (Jimmie) Savage. En 1954 obtiene una cátedra en la Facultad de Economía de la Universidad de Roma y en 1961 obtiene la de la Facultad de Ciencias de la misma universidad, en la que permanecerá hasta su jubilación en 1976. En 1970 aparece su opera magna, la *Teoria delle Probabilità* donde sintetiza sus ideas sobre probabilidad y estadística, y que ha sido traducida al inglés y al alemán. Poco después, a propuesta de Jimmie Savage, publica *Probability, Induction and Statistics* en 1972, que es una versión en inglés, -ampliada, corregida y reestructurada por el autor-, de importantes trabajos suyos originalmente publicados en italiano entre 1949 y 1967. Su conferencia invitada en el congreso del *International Statistical Institute* celebrado en Viena en 1973, constituye una lúcida descripción del papel unificador de los métodos bayesianos en los fundamentos y en las aplicaciones de la estadística. Tres años más tarde, ante la *First European Conference on New Developments and Applications of Bayesian Methods*, la primera conferencia internacional sobre métodos bayesianos que tuvo lugar en el mundo, celebrada en Fontainebleau en junio de 1976, de Finetti analiza el problema de las probabilidades de orden superior. Bruno de Finetti muere en Roma en 1985.

Un estudio sobre la vida y la contribución a la estadística de de Finetti, puede verse en Bernardo (1998). Casi toda la contribución de de Finetti puede verse en (de Finetti 1937, 1972, 1974 y 1975).

La contribución fundamental de de Finetti es, como ya se ha dicho, la noción de *intercambiabilidad*, noción más general que la de independencia y que permite justificar la aproximación bayesiana y los modelos jerárquicos, y su relación con los métodos frecuentistas.

La noción de muestra aleatorio simple, es ahora perfectamente natural dentro de la inferencia estadística y quiere decir que la distribución de la muestra puede escribirse como, con notación de función de densidad:

$$f(x_1, ..., x_n|\theta) = f(x_1|\theta) \cdots f(x_n|\theta)$$

.

Sin embargo, con esta definición, cuyo sentido es conseguir que las matemátics de la inferencia sean má simples, puede ocurrir que observemos repetidamente el mismo valor x_i, lo que obviamente es absurdo. La noción lógica a utilizar es la de observaciones *intercambiables*; la colección $(x_1, ...x_n)$ es intercambiable cuándo su distribución

$$f(x_1, ..., x_n) = f(x_{\pi(1)}, ..., x_{\pi(n)})$$

para cualquier permutación π de los n elementos. Es decir, la distribución de la muestra, sólo depende de los elementos que la componen. Es evidente que el concepto de *intercambiabilidad* es más débil que el concepto de *independencia*; lo más importante es que se demuestra, que, si la colección $(x_1, ...x_n)$ es intercambiable, la distribución conjunta se puede factorizar en la forma

$$f(x_1, ..., x_n) = \int_\Theta \Pi_{i=1}^n f(x_i|\theta)\pi(\theta)d\theta,$$

donde la distribución del parámetro $\pi(\theta)$ se puede obtener mediante límite de la distribución empírica de la muestra. Un desarrollo más preciso de estas ideas está contenido en *Bayesian Theory* (1994) página 177.

Este resultado justifica los modelos jerárquicos bayesianos donde las observaciones $x_i|\theta$ se comportan como variables independientes e idénticamente distribuidas, condicionalmente a θ, y el parámetro θ tiene una distribución inicial dada por $\pi(\theta)$.

La siguiente contribución a la aproximación bayesiana a la inferencia se debe a L. Savage (1917-1971) y otros, pero ya se puede decir que a partir de las contribuciones de Jeffreys y de Finetti la aproximación bayesiana ha recuperado la posición que nunca debió perder.

Agradecimientos

Este trabajo ha sido realizado en parte con ayudas del Ministerio de Ciencia e Innovación, proyecto MTM2008-03282 y de la Universidad Complutense de Madrid-BSCH, proyecto 910395.

BIBLIOGRAFIA

Bayes, T. (1764) An essay towards solving a problem in the doctrine of chances. Phillos. Trans. R. Soc. London, 53, 370-418, Reeditado por Deeming (1940) en Biometrika, 45, 293-315. Traducido al alemán con un comentario por Timerding (1908). Traducido al francés por Cléro (1988). Traducido al castellano por Gómez Villegas, et als. (2001) en *Rev. R. Acad. Cienc. Exact. Fis. Nat. (esp.)*, **95**, 1-2, 63-80.

Bernardo, J. M. (1998) Bruno de Finetti en la Estadística Contemporánea, En *Historia de la Matemática en el siglo XX*, Rev. Acad. Cienc. Exact. Fis. Nat., Madrid, 63-80.

Bernardo, J. M. and Smith, A. F. M. (1994) *Bayesian Theory*, London: John Wiley.

de Finetti, B. (1937) Foresight: its logical laws, its subjective sources. Reprinted in *Studies in Subjective Probability, 2nd edition*, 1980 (H. Kyburg and Smokler, eds.), 53-118, London: John Wiley.

de Finetti, B. (1972) *Probability, Induction and Statistics*, London: John Wiley.

de Finetti, B. (1974) *Theory of Probability. Volume 1*, London: John Wiley.

de Finetti, B. (1975) *Theory of Probability. Volume 2*, London: John Wiley.

Fisher, R. A. (1930) *The Genetical Theory of Natural Selection*, London: Oxford University Press, (Reprinted by Oxford University Press, Oxford, 1990).

Fisher, R. A. (1935) *The Design of Experiments*, Edinburg: Oliver & Boyd, (Reprinted by Oxford University Press, Oxford, 1990).

Fisher, R. A. (1956) *Statistical Methods and Scientific Inference*, Edinburg: Oliver & Boyd, (Reprinted by Oxford University Press, Oxford, 1990).

Girón, F. J. y Gómez Villegas, M. A. (1998) R. A. Fisher: su contribución a la ciencia estadística. Editado por la Real Academia de Ciencias Exactas, Físicas y Naturales en *Historia de la Matemática en el Siglo XX*, Madrid, 43-61.

Gómez Villegas, M. A. (2001) El ≪Ensayo Encaminado a resolver un problema en la Doctrina del Azar≫, *Rev. R. Acad. Cienc. Exact. Fisi. Nat. (esp.)*, **95**, 1-2, 81-85.

Gómez Villegas, M. A. (1994) El problema de la probabilidad inversa: Bayes y Laplace. Editado por E. Bustos y otros en Perspectivas Actuales de Lógica y Filosofía de la Ciencia, Ed. Siglo XXI: Madrid, 385-396.

Gómez Villegas, M. A. (2005) *Inferencia Estadística*, Madrid: Díaz de Santos.

Gómez Villegas, M. A., Girón, F. J., Martínez, M. L. y Rios, D. (2001) Un ensayo encaminado a resolver un problema en la doctrina de azar. *Rev. R. Acad. Cienc. Exact. Fisi. Nat.*, **95**, 1-2, 63-80.

Jeffreys, H. (1939, 1948, 1956) *Theory of Probability*, Oxford: Oxford University Press.

Laplace, P. S. (1774) Mémoire sur la probabilité des causes par les évènements. *Memoires de l'Academie Royale des Sciences Presents Par Divers Savans*, **6**, 621-656.

Laplace, P. S. (1812) *Thorie Analytique des probabilités*, Paris: Courcier.

Pearson, K. (1905) On the general theory of skew correlation and non linear regression. *Draper's Company Research Memoirs*, (Reprinted in Karl Pearson's Early Statistical Papers), Cambridge: Cambridge University Press),(first issued 1948), 477-528.

Stigler, S. M. (1986) *The History of Statistics: The Measurement of Uncertainty before 1900*, Cambridge: Belknap Harvard.