Rev.R.Acad.Cienc.Exact.Fis.Nat. (Esp),
Vol. 95, N.^{os} 1–2, pp 81–85, 2001
Monográfico: Conmemoraciones Matemáticas: Tricentenario de Thomas Bayes

EL «ENSAYO ENCAMINADO A RESOLVER UN PROBLEMA EN LA DOCTRINA DEL AZAR»

(Thomas Bayes/Ensayo de Bayes/Probabilidad Inversa)

MIGUEL ANGEL GÓMEZ VILLEGAS *
* Departamento de Estadística e I. O. Universidad Complutense de Madrid. Ciudad Universitaria s/n. 28040 Madrid.

ABSTRACT

This work starts with a biographical introduction including a few things known of Thomas Bayes life. It is followed by the background to the inverse probability problem and the comments about the Essay. Three applications added to the original job of Thomas Bayes by Richard Price are included. Several opinions about the Essay given by some scientific figures are embodied at the end of the paper.

RESUMEN

El trabajo consta de una introducción biográfica en el que se recogen las pocas cosas que se conocen de la vida de Thomas Bayes. A continuación se tratan los antecedentes del problema de la probabilidad inversa y se comenta el Ensayo. Se recogen también tres aplicaciones añadidas al trabajo original de Thomas Bayes por Richard Price. Se incluyen las opiniones que el Ensayo ha suscitado a algunos eminentes estadísticos.

1 NOTAS BIOGRÁFICAS

Lo primero que hay que decir al iniciar la biografía de Thomas Bayes es que pocas cosas son conocidas de la vida de nuestro autor. Ya en lo referente a su fecha de nacimiento Hald (1998), estadístico de la Universidad de Copenhague, y Stigler (1986), de la Universidad de Chicago, admiten que Bayes nació en el 1701, mientras que Maistrov (1974), el célebre historiador ruso continuador de la escuela de San Petersburgo, Karl Pearson (1978), el autor de las célebres conferencias sobre la Historia de la Estadística en los siglos XVI y XVII, Barnard (1958), quién se ocupó de que se restaurara por suscripción pública entre los estadísticos la tumba de Bayes, y Dale, el autor de la *Historia de la Probabilidad Inversa*, fijan el 1702 como la fecha de su nacimiento. Así con un razonamiento laplaciano, podríamos decir que Bayes nació con probabilidad 1/3 en el 1701 (en cuyo caso estaríamos ahora realmente en su tercer centenario) y con probabilidad 2/3 en el 1702.

Se sabe que en 1721 se traslada a Tunbridge Wells, inicialmente para ayudar a su padre que era ministro de la iglesia protestante, quedándose después en el mismo puesto. En 1731 escribe sobre *La Divina Benevolencia o un intento de demostrar que el principal fin de la Providencia divina es el gobierno y la felicidad de sus criaturas.*

Bajo el seudónimo de John Noon publica el trabajo titulado *An Introduction to the Doctrine of Fluxions and a Defence of Mathematicians against the Objections of the Author of the Analyst* donde defiende a los matemáticos aplicados de las objeciones que planteaba el obispo Berkeley, autor de *El Analista*, al nuevo cálculo diferencial de Newton. Como se ve, la polémica entre matemática pura y aplicada viene de antiguo.

En 1742 Bayes es elegido miembro de la Royal Society. Por cierto que se ha especulado sobre el que fuera elegido teniendo tan escasas contribuciones científicas. La propuesta para su elección dice:

> *El Reverendo Thomas Bayes de Tunbridge Wells, deseando el honor de ser elegido miembro de la Sociedad, es propuesto y recomendado como un caballero de méritos conocidos, bien preparado en geometría y en todas las partes de la matemática y de la filosofía,*

y cualificado en todos los aspectos para ser un valioso miembro de la Sociedad.

Apoyan la propuesta con su firma Sthanhope, el conde de Chesterfield, cuyas cartas a su hijo natural constituyen un clásico de la literatura inglesa; Martin Folkes, el presidente de la Royal Society en aquel momento; Cromwell, médico y descendiente del lord protector, James Burrow quien sería presidente de la Royal Society entre 1768 y 1772; y John Eames tutor de una academia no-conformista y quizás la única persona del círculo de Bayes entre los firmantes. Sin duda todos ellos constituyen un plantel de personalidades con las que el nombramiento no podía serle negado.

Bayes se retira de su ministerio en 1752. El 17 de Abril de 1761 muere en Tunbridge Wells y es enterrado en *Bunhill fields* el mismo cementerio en el que están enterradas otras figuras importantes como Price, Defoe, etc. de la corriente no-conformista.

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, amigo de Bayes y también miembro de la Royal Society. Se trata sin duda de uno de los trabajos que más discusiones ha suscitado en el campo de la estadística. En este momento Laplace tiene 15 años, siendo este dato importante porque durante cierto tiempo se consideró la posibilidad de que Laplace plagiara este trabajo; hoy en día está fuera de duda el que el teorema de Bayes fue redescubierto por Laplace sin inspirarse en el Ensayo.

La otra contribución a la matemática que conocemos de Bayes se debe tambien a Price y es una nota en la que se prueba la divergencia de la serie $ln(z!)$.

2 ANTECEDENTES

El antecedente más directo que podemos citar del problema 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 en n repeticiones. El problema es recogido por De Moivre (1667-1754), la mayor autoridad del momento en el emergente cálculo de probabilidades. Este hugonote francés que se ve obligado a exiliarse a Inglaterra, trata el problema en su libro *Doctrine of Chances*. Con notación actual podemos 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.

Pero tanto Bernoulli como De Moivre hacen aseveraciones acerca de *problemas directos* de probabilidad, en los que dando por conocida la probabilidad de éxito θ, se calcula la probabilidad de cualquier sucesión de éxitos y fracasos; por ejemplo la probabilidad de EEFEF se hace mediante la expresión

$$P\{EEFEF\} = \theta^3(1-\theta)^2$$

El problema que es abordado por Bayes en su Ensayo es lo que se llamó desde entonces el *problema inverso*. Se trata realmente de un problema de inferencia, en el que se pretende conocer la causa θ por su efecto r. En su 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 introducir una probabilidad inicial sobre θ y definir la densidad condicional $\pi(\theta\,|\,r)$. Desde luego es notable que Bayes diera estos dos pasos y pone de manifiesto la alta capacidad matemática del mismo. El descubrimiento fue hecho otra vez unos años más tarde por Laplace (1794-1827) quien en su *Theorie Analitique des Probabilites* (1812), obtiene además la distribución final para modelos más generales que el binomial.

3 EL ENSAYO

El escrito *An Essay Towards Solving a Problem in the doctrine of Chances* está compuesto por cuatro partes: una carta preliminar de envío por Price a la Royal Society, el enunciado del problema, dos secciones y un apéndice.

Parece que los papeles originales fueron encontrados por los deudos de Bayes y entregados por éstos a Price para que juzgara sobre su interés. Éste los estudia y se da cuenta del alto valor matemático de los mismos lo que le hace reescribir el Ensayo añadiendo algunas cosas, para después comunicárselo a la Academia. En la carta, Price dice que Bayes había escrito una introducción, que no nos ha llegado, y afirma haber añadido un escolio en la sección segunda así como las aplicaciones del apéndice.

El enunciado del problema está expuesto con claridad meridiana:

Dado el número de veces que un suceso ha ocurrido o fallado calcular la probabilidad de que la probabilidad de su ocurrencia en un solo experimento esté entre dos valores de probabilidad conocidos.

Se pretende por tanto calcular la probabilidad

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

Sigue la Sección I, que contiene 7 proposiciones, 5 corolarios y 2 definiciones, y acaba con la obtención de la distribución binomial. En esta primera sección se incluye la siguiente definición de probabilidad:

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.

Lo único que parece claro en esta definición es su oscuridad y que es una definición de probabilidad subjetiva.

La proposición 2 permite clarificar el significado que para Bayes tenía la probabilidad; dice así

Si una persona tiene una esperanza que depende de la ocurrencia de un suceso, la probabilidad del suceso es a la probabilidad de su fallo como lo que pierde si el suceso no ocurre es a su ganancia si el suceso ocurre.

Es decir para definir $P(A)$ se fija un valor arbitrario N y se pide la determinación subjetiva de una cantidad a que debería pagar uno, por que le dejen jugar al siguiente juego: si ocurre el suceso A se recibe la cantidad N y si no ocurre A (ocurre $\overline{A}$, el complementario de A) se pierde la cantidad a (se gana -a); para que el juego sea equitativo ha de cumplirse que

$$P(A)(N-a)+P(\overline{A})(-a)=0$$

de donde

$$P(A)=a/N.$$

Mediante este mecanismo, se puede definir la probabilidad del suceso A y comprobar que se siguen con todo rigor todas las aseveraciones sobre la probabilidad que se hacen en la primera sección del Ensayo (ver Gómez Villegas (1994)). Bayes se anticipa, por tanto, a la definición subjetivista de la probabilidad.

La Sección II contiene 1 postulado doble, 2 lemas, 3 proposiciones, 1 corolario, 1 escolio, 5 puntos y 3 reglas. En ella obtiene la distribución final.

Es interesante el postulado doble donde introduce el simil geométrico de un experimento de Bernoulli, mediante el lanzamiento sobre una mesa de billar de una bola:

1. *Yo supongo la tabla cuadrada ABCD del plano y que dos bolas O y W son lanzadas sobre ella; hay la misma probabilidad de que se detengan en una parte o en otra de la tabla.*

2. *Yo supongo que la bola W ha sido arrojada primero, y que por el punto por donde se detenga se traza una línea paralela al lado AD, que se encuentra con CD y AB en s y o respectivamente y que depués la bola O es arrojada n veces y que queda entre AD y os una de las veces, entonces digo que el suceso M ha ocurrido en un experimento.*

Obsérvese que si los lados de la tabla son unitarios y el segmento OA mide θ la probabilidad de que la bola quede entre AD y os es θ.

La proposición 8 de la Sección II contiene la expresión de la probabilidad conjunta de las variables θ y X; en notación actual

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

En el corolario que sigue a la proposición 8 se obtiene la distribución marginal

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

La proposición 9 contiene el cálculo de la probabilidad final pedida; en notación actual

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

Es la versión continua del teorema de Bayes. Si se supone como distribución inicial la uniforme en el intervalo $[0,1]$ y el modelo binomial $P\{X=r\,|\,\theta\}$, mediante las expresiones

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

calcula la densidad final

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

Después de recoger la expresión de la distribución final, incluye como corolario la fórmula

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

que es la particularización de la proposición 9 al intervalo $(b, 1)$.

A continuación se incluye el célebre escolio en el que justifica la utilización de la distribución inicial uniforme, y del que Price habla en la carta de envío del ensayo a la Royal. Este escolio o justificación es debido a Price y ha sido visto por algunos como el germen de la aproximación a la inferencia predictiva, traducimos textualmente:

> *...en lo que concierne al suceso, no hay razón para pensar que en un cierto número de experimentos debería ocurrir cualquiera de los posibles resultados un mayor número de veces que otro...*

En nuestra opinión el significado del escolio, salvo lo que se acaba de traducir, no es muy claro y parece adivinarse en él una justificación para la utilización de la distribución inicial uniforme. De todas formas a la luz del corolario de la proposición 8 cabe suponer que Bayes se diera cuenta de que con las suposiciones hechas se tiene que

$$\begin{aligned} P\{X = j\} &= \int_0^1 P\{X = j \,|\, \theta\}\pi(\theta)\, d\theta \\ &= \binom{n}{j} B(j+1, n-j+1) \\ &= \frac{1}{n+1} \end{aligned}$$

para todo $j = 0, \ldots, n$, lo que de hecho apoya la utilización de la distribución uniforme, ya que asegura que la probabilidad de observar en n repeticiones el suceso un número de veces 0,1,2 o n es siempre la misma e igual a $1/(n+1)$.

La sección II acaba con la proposición 10, 5 puntos y 3 reglas. Todos estos elementos son aproximaciones para calcular la función beta incompleta que aparece en la proposición 9, y que permiten el cálculo de la distribución final.

4 TRES APLICACIONES

El Ensayo termina con un Apéndice que contiene la aplicación de la probabilidad final a tres casos particulares. Este apéndice es todo él debido a Price.

La primera aplicación incluye un cálculo exacto de la probabilidad final, en concreto obtiene la probabilidad de que un suceso se presente en la segunda realización del experimento, si se ha presentado ya la primera vez. En la segunda calcula la probabilidad de que amanezca al día siguiente, en concreto enuncia: una persona que llegara a la Tierra y viera amanecer, para calcular la probabilidad de que amaneciera al día siguiente, debiera calcular

$$P\{X_2 = 1 \,|\, X_1 = 1\} = \frac{P\{X_1 = 1, X_2 = 1\}}{P\{X_1 = 1\}} = \frac{2}{3};$$

también obtiene la expresión

$$\begin{aligned} P\{X_p = 1 \,|\, X_1 = 1, \ldots, X_{p-1} = 1\} &= \\ \frac{P\{X_1 = 1, \ldots, X_p = 1\}}{P\{X_1 = 1, \ldots, X_{p-1} = 1\}} &= \frac{p}{p+1} \end{aligned}$$

y observa que su límite es 1 cuando p crece, aunque añade que la certeza de las cosas se obtiene viendo la ley que permite su ocurrencia, no observando el límite de las probabilidades.

La tercera aplicación es para estimar la proporción de premios existentes en una lotería. Determina la estimación en términos de apuesta que debe de hacer una persona que sin conocer nada del mecanismo oiga que han salido 10 boletos sin premio y uno premiado.

5 COMENTARIOS

Como se ha dicho el Ensayo es una de las contribuciones estadísticas que más polémica ha suscitado. Laplace (1812) opina que es ciertamente oscuro, este autor, que repasaba varias veces sus propios escritos, debió encontrar ciertamente enmarañado el método geométrico de demostración de Bayes. Todhunter (1865), el autor de *La Historia de la Teoría Matemática de la Probabilidad*, dice que el resumen que antecede en el Ensayo a la proposición 9 es algo oscuro. Molina (1930) sostiene que Bayes y Price pueden ser clasificados como grandes matemáticos entre los que les precedieron y siguieron. K. Pearson (1978), en sus celebradas conferencias sobre historia de la probabilidad recogidas por su hijo E. Pearson sostiene que Bayes y Price son totalmente conscientes de su novedosa resolución del problema opuesto al de De Moivre. Fisher afirma que la contribución de Bayes coloca a éste en el primer orden de los pensadores independientes. Barnard, ya citado, escribe que el trabajo matemático de Bayes es de la más alta calidad. Por último Stigler (1999), en nuestra opinión el estadístico-historiador que ha tratado con más profundidad estos temas, dice del Ensayo que es extremadamente difícil de leer, pero que es fundamental su contribución al problema inverso y anticipa la aproximación predictiva a la inferencia.

Estas opiniones finales de figuras insignes de la estadística queremos que hoy constituyan nuestra humilde contribución al probable tercer centenario del nacimiento de Thomas Bayes.

AGRADECIMIENTOS

Esta investigación ha sido financiada en parte por el Ministerio de Educación y Cultura DGESIC PB98-0797. Quiero agradecer también a mi colega Javier Girón su amabilidad al recoger mi idea sobre la celebración de este centenario y su trabajo de homogeneizar las distintas contribuciones. Muchas gracias a los profesores Eusebio Gómez y Juan-Miguel Marín por haberme corregido algunos errores.

BIBLIOGRAFÍA

1. Barnard, G. A. (1958) *Thomas Bayes-a biographical note*, (thogether with a reprinting of Bayes, 1764), *Biometrika*, **45**, 293-315.
2. Bernoulli, J. (1713) *Ars Conjectandy*, Basilea: Thurnisiorum.
3. De Moivre, A. (1738) *The Doctrine of Chances*, Reprinted 1967, New York: Chelsea.
4. Maistrov, L. E. (1974) *Probability Theory-A Historical Sketh*, New York: Academic Press.
5. Gómez Villegas, M. A. (1994) *El problema de la probabilidad inversa: Bayes y Laplace*, (Editado por, E. de Bustos, y otros), En *Perspectivas Actuales de Lógica y Filosofía de la Ciencia*, Madrid: Siglo XXI, 385-396.
6. Hald, A. (1998) *A History of Mathematical Statistics from 1750 to 1930*, New York: Wiley.
7. Laplace, P. S. (1812) *Théorie Analytique des Probabilités*, Paris: Courcier.
8. Molina, E. C. (1930) The theory of probability: some comments on Laplace's théorie analytique. *Bulletin of the American Mathematical Society*, **36**, 369-392.
9. Pearson, K. (1978) *The History of Statistics in the 17th and 18th Centuries, against the Changing Background of Intellectual, Scientific and Religious Thought*, lectures from 1921-1933, Ed. E. S. Pearson, London: Griffin.
10. Stigler, S. M. (1999) *Statistics on the Table the History of Statistical Concepts and Methods*, Cambridge: Harvard University Press.
11. Thodhunter, I. (1865) *A History of the Mathematical Theory of Probability from the Time of Pascal to that Laplace*, London: MacMillan. Reprinted, 1949, 1965, New York: Chelsea.