Ensayos Bayesianos I: Buscando a Bayes

MAURICIO CANALS LAMBARRI*

PROGRAMA DE SALUD AMBIENTAL, ESCUELA DE SALUD PÚBLICA
FACULTAD DE MEDICINA, UNIVERSIDAD DE CHILE

Resumen

En este ensayo revisamos las ideas originales de Thomas Bayes y Pierre Simon Laplace, buscando las bases de lo que conocemos hoy como teorema de Bayes. Vemos que la expresión actual de éste se fue construyendo entre mediados del siglo XVIII e inicios del siglo XX y tiene poco que ver con los planteamientos iniciales. Mostramos la importancia de Laplace en su desarrollo y algunas controversias relativas a su fondo y forma de expresión.

Introducción

Las formas como se ha enunciado el llamado teorema de Bayes han variado en el tiempo [1]. Hoy lo habitual es que llamemos teorema de Bayes a una expresión que se fue estableciendo paulatinamente en el tiempo desde mediados del siglo XVIII hasta principios del siglo XX. Tal vez la referencia más antigua en que aparece la expresión actual bajo el nombre de *Formule de Bayes* se encuentra en *Calcul des Probabilites* de Henri Poincaré en 1912 [2]. Habitualmente enunciamos como teorema de Bayes (1763) la siguiente expresión: "Dada una partición del espacio muestral C_i y un evento A, entonces la probabilidad condicional (o *a posteriori*) de C_i, dado el evento A, es:

$$\mathrm{P}(C_i \mid A) = \frac{\mathrm{P}(A \mid C_i)\mathrm{P}(C_i)}{\sum\limits_j \mathrm{P}(A \mid C_j)\mathrm{P}(C_j)} \tag{1}$$

Podemos demostrar este teorema hoy en forma muy sencilla. Primero debemos conocer la definición (o axioma) de probabilidad condicional: "la probabilidad de un evento A cuando ha ocurrido un evento B ($\mathrm{P}(A \mid B)$) es $\mathrm{P}(A \mid B) = \mathrm{P}(A \cap B)/\mathrm{P}(B)$, o equivalentemente $\mathrm{P}(A \cap B) = \mathrm{P}(A \mid B)\mathrm{P}(B)$".

A continuación podemos demostrar el teorema de la probabilidad total, el que enunciamos de la siguiente forma: "dada una partición del espacio muestral C_i y un evento A, entonces la probabilidad total del evento A es: $\mathrm{P}(A) = \sum_j \mathrm{P}(A \mid C_j)\mathrm{P}(C_j)$". Es fácil ver que $A = \cup_j (A \cap C_j)$ y como los C_j son independientes (y además excluyentes), usando la definición de probabilidad condicional tenemos

$$\mathrm{P}(A) = \sum_j \mathrm{P}(A \cap C_j) = \sum_j \mathrm{P}(A \mid C_j)\mathrm{P}(C_j) \tag{2}$$

Después de esta demostración, el teorema de Bayes se vuelve trivial:

$$\mathrm{P}(C_i \mid A) = \frac{\mathrm{P}(C_i \cap A)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A \cap C_i)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A \mid C_i)\mathrm{P}(C_i)}{\mathrm{P}(A)}$$

y por el teorema de la probabilidad total se obtiene finalmente

$$\mathrm{P}(C_i \mid A) = \frac{\mathrm{P}(A \mid C_i)\mathrm{P}(C_i)}{\sum\limits_j \mathrm{P}(A \mid C_j)\mathrm{P}(C_j)} \tag{3}$$

Sin embargo, cuando acudimos a las fuentes originales de este teorema, todo se vuelve obscuro y nebuloso (Figura 1). De partida, se ha cuestionado que la imagen conocida de Bayes, sea realmente la de él e incluso su fecha de nacimiento [3]. Por otra parte, el teorema supuestamente se formula en 1763 y Bayes murió en 1761, y cuando se le da una primera lectura al artículo original, este teorema no es evidente en ninguna parte.

El objetivo de este artículo es tratar de seguir el origen histórico y analizar los artículos fundamentales hasta llegar a lo que hoy conocemos como teorema de Bayes.

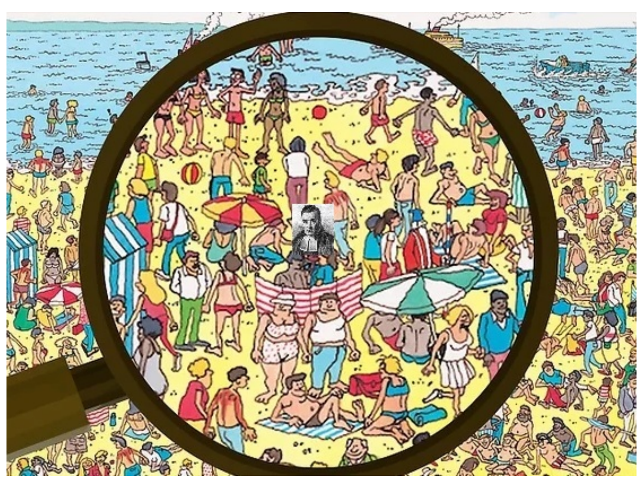

Figura 1: ¿Donde está Bayes? Encontrar el teorema de Bayes en el ensayo original de 1763 es tarea difícil.

*mcanals@uchile.cl

Análisis

Thomas Bayes (1701-1761) fue un matemático y un reverendo presbiteriano. Perteneció a la *Royal Society of London* desde 1742, probablemente por su trabajo "*An introduction to the doctrine of fluxions, and a defence of the mathematicians against the objections of the author of the analyst*", que fue publicada en forma anónima en 1736, en la que defendía la ley de gravitación de Newton. Sin embargo, según Richard Price (1723-1791), Bayes estaba interesado en estudiar las causas más probables de aquello que sucedía para poder probar la existencia de Dios y su benevolencia [4]. No obstante esto es una interpretación de Price, quien era su amigo. No se conoce otra publicación matemática de Bayes. Richard Price, destacado clérigo y escritor británico, recibió como legado a la muerte de Bayes cien libras y un conjunto de artículos sin terminar, en los que probablemente trabajó el mismo [5]. En 1763 Price publicó en una comunicación a John Canton el trabajo de Bayes [4]. En esta publicación dio a conocer la obra de Bayes, que "contendría" el teorema.

Sin embargo, al leer este artículo hay varios aspectos que merecen duda o discusión. El primero es la real motivación de Bayes al escribirlo. De hecho Price lo presenta como una respuesta al segundo libro de David Hume (1748), en que se manifiesta abiertamente escéptico de los milagros. Pero se ha discutido la posibilidad que en realidad sea un estudio del problema de la "probabilidad inversa" iniciado en la segunda edición de la "Doctrina de las Chances" de Abraham De Moivre (1667-1754) en 1718 [6]. El segundo es ¿cuál fue el título original del trabajo? En *Philosophical Transactions* aparece como "*An essay towards solving a problem in the doctrine of chances* (1763), pero en las *separatas* del trabajo, aparece como "*A method of calculating the exact probability of all conclusions founded on induction* (1764), un título que representa mucho más el contenido del artículo (ver discusión en [5]). El tercer problema es que al leer el trabajo, bajo la nebulosa, aparecen visos del teorema.

Analicemos algunas partes del trabajo original de 1763.

<u>Sección 1; Proposición 5.</u> "*If there be two subsequent events, the probability of the 2nd b/N and the probability of both together P/N and it being 1st discovered that the 2nd has happened, from hence I guess that the first event has also happened, the probability I am in the right is P/b*".

En esta proposición y usando la proposición 3 del mismo ensayo, que se refiere a la probabilidad que el primer evento (A) y el segundo evento (B) ocurran ($A \cap B$), Bayes nos propone que si $\mathrm{P}(B) = b/N$ y $\mathrm{P}(A \cap B) = P/N$, entonces $\mathrm{P}(A \mid B) = P/b$. En otras palabras $\mathrm{P}(A \mid B) = \mathrm{P}(A \cap B)/\mathrm{P}(B) = (P/N)/(b/N) = P/b$. Así, en las proposiciones 3 y 5 Bayes construye la definición de probabilidad condicional y la relaciona con la probabilidad de la ocurrencia conjunta (intersección) de dos eventos, lo que hoy es considerada una definición o bien un axioma [7]. Sin embargo, aún no es el teorema.

<u>Sección 1; Proposición 7.</u> Aquí Bayes construye la distribución binomial a partir de la repetición de ensayos de Bernoulli, pues la usará después en el ensayo. Aquí hay que tener cuidado con la notación de Bayes, que hoy nos complica. Bayes utiliza las letras a y b para las probabilidades y p y q para las réplicas, justo a la inversa de lo que hacemos hoy.

"*If the probability of an event be a, and that of his failure be b in each single trial, the probability of its happening p times, and failing q times in $p+q$ trials is Ea^pb^q, if E be the coefficient of the term in which occurs a^pb^q when the binomial $(a+b)^{p+q}$ is expanded.*".

En palabras de hoy: $\mathrm{P}(X = p) = \binom{p+q}{p} a^p b^q$.

<u>Sección II.</u> En realidad Bayes no se preocupó sustancialmente de estos resultados, pero los incluyó para responder un problema más general que se refiere a encontrar la probabilidad de ocurrencia de un evento cuando este ya ha ocurrido y fallado antes cierto número de ocasiones: "*Given the number of times in which an unknown event has happened and failed [... Find] the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named.*".

En esta sección Bayes se imagina una "mesa de billar", un rectángulo ABCD sobre el cual se lanzan bolas al azar. Interesantemente Bayes una vez más es complicado y parece hacer todo al revés. Nomina las letras del rectángulo en sentido matemático negativo (sentido de las manijas del reloj) y mide sus distancias de derecha a izquierda (Figura 2).

<u>Lema 1.</u> Simplificando a Bayes. Si se tira una bola W al azar y se traza una línea vertical "os" sobre el punto en que esta cae (Figura 2), entonces la probabilidad que caiga entre cualesquiera dos puntos (f, b) es la razón entre la distancia entre los puntos y la distancia AB. Hoy la demostración sería muy simple. Por teoría de la medida la probabilidad es el cociente entre la medida del evento y la medida del espacio muestral: $\mathrm{P} = m(A)/m(\Omega)$. Si todos los puntos en que puede caer la bola W <u>son igualmente probables</u>, entonces una medida del evento son todos los puntos del rectángulo $bLFf$, es decir su área $= d(f, b) \times AD$ y una medida del espacio muestral son todos los puntos de la mesa $ABCA$, es decir su área $= AB \times AD = ABAD$. Entonces la probabilidad buscada $\mathrm{P}(x \in (f, b)) = d(f, b)/AB$.

Lema 2. Aquí Bayes denomina "evento M" a "la bola W cae entre la recta os y el borde de la mesa AD", y enuncia: "*The ball W having been thrown, and the line os drawn, the probability of the event M in a single trial is the ratio of Ao/AB*", lo cual es sólo un corolario del lema anterior: $\mathrm{P}(M_1) = d(Ao)/AB$.

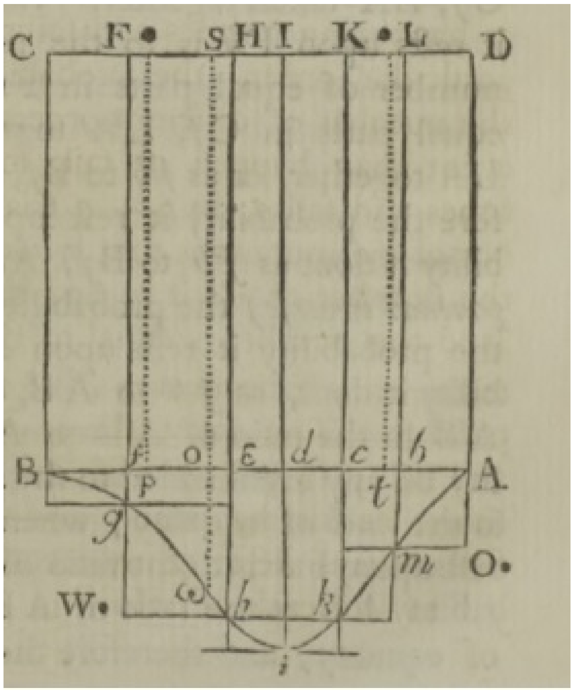

Figura 2: Figura original del ensayo publicado en 1763, donde se reconoce la "mesa" $ABCD$, la línea os que determina la ocurrencia del evento M y la construcción de Bayes de la curva binomial dependiendo de la posición relativa del punto en que cae la bola.

Y ahora, una vez preparada la cancha, nuestro reverendo lanza su primer misil.

Proposición 8. Para entender esta proposición hay que mirar la Figura 2, obtenida del original de Bayes. "*I say that before the ball is thrown, the probability the point o should fall between f and b, any two ponits nemed in the line AB, and withal that the event M should happen p times and fail q in $p+q$ trials, is the ratio of $fghikmb$, the part of the figure $BghikmA$ intercepted between the perpendiculars fg, bm raised upon the line AB, to CA the square upon AB*".

Hoy lo simplificaríamos diciendo que la probabilidad buscada es el área bajo la "curva binomial" entre los puntos f y b partido por el área total de la "mesa de billar". Es importante notar que aquí Bayes quiere calcular la probabilidad de la intersección entre el evento $X \in [f, b]$, siendo X la coordenada de un punto al azar sobre la línea AB determinado por la caída de la bola W, y el evento M_p^{p+q}, es decir, se tienen p caídas de la bola W en el intervalo determinado por la recta os y AD, en $p+q$ ensayos (ver Figura 3). Notamos que si se relativiza la posición X con respecto a AB, entonces $x = X/AB$ varía entre 0 y 1 y además x representa la probabilidad del evento M para una posición cualquiera X. Aunque no lo expresa en esta forma, Bayes nos dice que la probabilidad buscada es:

$$\mathrm{P}\left(X \in (f, b) \cap M_p^{p+q}\right) = \int_f^b \binom{p+q}{p} x^p (1-x)^q dx \tag{4}$$

Según Timerding (1908) "Para entender la presentación de Bayes, hay que recordar que en Inglaterra la designación integral estaba mal vista porque su creador, Leibniz, era considerado un plagiario de Newton" (en [1]).

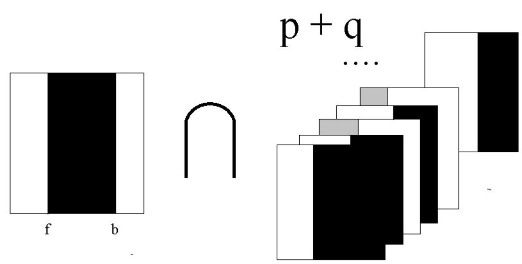

Figura 3: La proposición 8 de Bayes se trata de la ocurrencia de los eventos "la bola cae entre los puntos f y b" y "ocurre el evento M p veces en $p+q$ intentos".

Entonces Bayes lanza su segundo misil, aún más poderoso.

Proposición 9. "*If before anything is discovered concerning the place of the point o, it should appear that the event M had happened p times and failed q in $p+q$ trials, and from hence I guess that the point o lies between any two points in the line AB, as f and b, and consequently that the probability of the event M in a finite trial was somewhere between the ratio of AB to AB and Af to AB: the probability I am in the right is the ratio of that part of the figure A_iB described as before which is intercepted between perpendiculars erected upon AB at the points f and b, to the whole figure A_iB*".

Decifrando esta "retórica", Bayes nos dice:

$$\mathrm{P}\left(X \in (f, b) \mid M_p^{p+q}\right) = \frac{\int_f^b \binom{p+q}{p} x^p (1-x)^q dx}{\int_0^1 \binom{p+q}{p} x^p (1-x)^q dx}, \tag{5}$$

lo cual se sigue casi de inmediato de la Sección I, proposición 5 y la proposición 8. Solo basta ver que

$$\begin{aligned} &\mathrm{P}\left(X \in (0,1) \cap M_p^{p+q}\right) \\ &\quad = \mathrm{P}\left(M_p^{p+q} \mid X \in (0,1)\right) \mathrm{P}\left(X \in (0,1)\right) \\ &\quad = \int_0^1 \binom{p+q}{p} x^p (1-x)^q dx, \end{aligned} \tag{6}$$

basados en que cada punto del intervalo que señala la abscisa del lugar en que cae la bola W, tiene igual probabilidad de ocurrir (distribución *a priori* uniforme).

Y en su "escolio" agrega una importante frase: "Por lo tanto, en lo que sigue daré por sentado que la regla dada con respecto al evento M en la proposición 9 es también la regla que debe usarse en relación con cualquier evento respecto de cuya probabilidad no se sabe nada en absoluto con anterioridad a cualquier ensayo realizado u observado al respecto. Y a tal evento lo llamaré un evento desconocido". Debemos notar que ésta es la regla de Bayes y que contiene un argumento de una distribución *a priori* uniforme [8].

En la siguiente proposición nuestro reverendo deja volar su imaginación y nos dice:

Proposición 10. "*I say then that if an unknown event has happened p times and failed q in $p+q$ trials, and in the base taking any points as f and t, the chance that the probability of the event lies somewhere between the ratio Af/AB and At/AB is the part of therefore figure which intercepted between the two ordinates, to the whole figure on the base AB*". Aquí he modificado levemente la redacción de Bayes ya que obligaba a hacer una nueva figura, cambiar las letras y enredar todo.

Observamos que ahora no se trata de un punto físico donde cae una bola sino en qué intervalo cae una probabilidad (w) de un evento de cuya probabilidad no se sabe nada (*a priori* uniforme), pero que ha ocurrido p veces en $p+q$ ensayos. En otras palabras ahora trata al parámetro w (una probabilidad) como una variable aleatoria que puede caer en un intervalo.

$$\mathrm{P}\left(w \in (f,t) \mid M_p^{p+q}\right) = \frac{\int_f^t \binom{p+q}{p} w^p (1-w)^q dw}{\int_0^1 \binom{p+q}{p} w^p (1-w)^q dw} \tag{7}$$

La demostración es obvia por la propia construcción de Bayes al asignar las abscisas como razones, que a su vez representan probabilidades.

El resto del ensayo Bayes-Price son ejemplos y agregados de Price, pero que no añaden más a la teorización. Sin embargo hay que destacar que aunque Bayes no menciona en ninguna parte la probabilidad de las causas de un evento determinado pero Price si ve en el ensayo de Bayes una aplicación en ese sentido, haciéndolo explícito en su "*covering letter*" acompañando el artículo, dirigido a John Canton y además en su apéndice al ensayo de Bayes [1]. Posteriormente Pierre Simon Laplace (1749-1827) aborda directa e intencionadamente este problema:

> "*je me propose de déterminer la probabilité des causes par les événements, matière neuve à bien des égards et qui mérite d'autant plus d'être cultivée que c'est principalement sous ce point de vue que la science des hasards peut être utile dans la vie civile*"

Laplace (1774) presentó sus cálculos en forma independiente del teorema de Bayes, pero explícitamente en relación a las causas de un evento [9]:

> "*Si un événement peut etre produit par un nombre n de causes diferentes, les probabilités de léxistence de ces causes prises de l´énvénement sont entre elles comme les probabilités de l´énvénement prises de ces causes et la probabilité de l´existence de chacune d´elles est egale a la probabilité de l´énvénement prise de celle cause divisse par la somme de toutes les probabilités de l´énvénement prises de chacune de ces causes*"

(Si un evento puede ser producido por un número n de causas diferentes, las probabilidades de existencia de estas causas dado el evento son como las probabilidades del evento dado estas causas y la probabilidad de existencia de cada una de ellas es igual a la probabilidad del evento tomado de esa causa dividida por la suma de todas las probabilidades del evento tomado de cada una de estas causas).

En nuestros días esto se puede expresar como:

$$\frac{\mathrm{P}(C_i \mid A)}{\mathrm{P}(C_j \mid A)} = \frac{\mathrm{P}(A \mid C_i)}{\mathrm{P}(A \mid C_j)} \tag{8}$$

$$\mathrm{P}(C_i \mid E) = \frac{\mathrm{P}(E \mid C_i)}{\sum_j \mathrm{P}(E \mid C_j)} \tag{9}$$

Como vemos, esta expresión es el teorema de Bayes, también para una distribución *a priori* uniforme, ¡pero expresado para una variable aleatoria discreta!, lo que es más parecido a nuestra notación actual. Notemos que Laplace no menciona a Bayes y al parecer sus resultados fueron completamente independientes.

Laplace plantea el siguiente problema: "Si una urna contiene una infinidad de fichas blancas y negras en una proporción desconocida y se sabe que en una tirada en que se sacan $m+n$ fichas hay n blancas (y m negras) (B_n^{m+n}). ¿Cuál es la probabilidad que en una nueva tirada, sacando una ficha, ésta salga blanca (B_1^1)?". Es decir, Laplace buscaba $\mathrm{P}(B_1^1/B_n^{m+n})$. Es claro que si en una urna la proporción de blancas es p, entonces, en una "sacada" la probabilidad de sacar una ficha blanca es también p. Laplace entonces imaginó infinitas urnas (o causas, C_i), cada una con una diferente proporción de blancas $p \in [0,1]$, todas equiprobables (*a priori* uniforme) y pensó que la probabilidad de los eventos va a tener relación con la urna de la cual proviene.

Así,

$$\begin{aligned} \mathrm{P}\left(C_i \mid B_n^{m+n}\right) &= \frac{\mathrm{P}\left(B_n^{m+n} \mid C_i\right)\mathrm{P}(C_i)}{\int \mathrm{P}\left(B_n^{m+n} \mid C_i\right)\mathrm{P}(C_i)} \\ &= \frac{\mathrm{P}\left(B_n^{m+n} \mid C_i\right)}{\mathrm{P}\left(B_n^{m+n}\right)}, \end{aligned} \tag{10}$$

y como cada repetición es un ensayo de Bernoulli, llamando x a la proporción desconocida de blancas en la urna tenemos

$$\mathrm{P}\left(x \mid B_n^{m+n}\right) = \frac{\binom{m+n}{n} x^n (1-x)^m dx}{\int_0^1 \binom{m+n}{n} x^n (1-x)^m dx}, \tag{11}$$

con el dx en el denominador dando cuenta de que al ser x una variable continua (al estar asociada a infinitas urnas), la probabilidad en un punto es 0.

Ahora la probabilidad buscada es igual a

$$\begin{aligned} \mathrm{P}\left(B_1^1 \mid B_n^{m+n}\right) &= \frac{\mathrm{P}\left(B_1^1 \cap B_n^{m+n}\right)}{\mathrm{P}\left(B_m^{m+n}\right)} \\ &= \frac{\int_0^1 \binom{m+n}{n} x^{n+1} (1-x)^m dx}{\int_0^1 \binom{m+n}{n} x^n (1-x)^m dx} \\ &= \frac{\int_0^1 x^{n+1} (1-x)^m dx}{\int_0^1 x^n (1-x)^m dx} \end{aligned} \tag{12}$$

Este resultado los extendió al caso de estimar la probabilidad $\mathrm{P}\left(B_r^{r+s} \mid B_n^{m+n}\right)$, obteniendo

$$\mathrm{P}\left(B_r^{r+s} \mid B_n^{m+n}\right) = \frac{\int_0^1 x^{n+r} (1-x)^{m+s} dx}{\int_0^1 x^n (1-x^m dx)} \tag{13}$$

Marie Jean Antoine Nicolas de Caritat, marqués de Condorcet (1743-1794), Francia, fue un matemático brillante y, al igual que Laplace, discípulo de D´Alambert, que además fue político, filósofo y científico. Sus escritos probabilísticos han sido poco estudiados y poco se sabe de las fechas exactas de cuando fueron escritos. Sin embargo, es claro que arribó a los mismos resultados que Bayes y Laplace entre 1777 y 1780 [10]. Al menos en 1777 escribe la fórmula de Laplace-Bayes como una regla ya establecida, sin demostrarla y en su manuscrito 873, problema 1, de fecha desconocida, obtiene la regla de Bayes. Es interesante que Karl Pearson llama a lo que él considera como teorema de Bayes, la extensión de Condorcet-Laplace de la regla de Bayes [11].

Como vemos, Bayes y Laplace llegan a resultados convergentes con ideas diferentes. Bayes al parecer pensando más en el método de inducción y Laplace más en la determinación de la probabilidad de las causas, entendidas éstas como discretas, lo que ya había vislumbrado Price. Ambos usan distribuciones *a priori* uniformes. Bayes para "los puntos" en los que cae la bola W y Laplace para sus infinitas urnas. Sin embargo, Laplace el 1778 eliminó la restricción de una distribución *a priori* uniforme [1], conduciendo a la formulación actual del teorema de Bayes. Buscando la expresión actual del teorema, vemos que en Poincaré (1916) ya se encuentra. En su libro *Calcul des probabilités* muestra que si C_i son las posibles causas de un evento A, cuyas probabilidades *a priori* son $\mathrm{P}(C_i) = w_i$ y las probabilidades $\mathrm{P}(A \mid C_i) = p_i$, entonces $\mathrm{P}(C_i \mid A) = p_i w_i / \sum_j p_j w_j$. Para Pearson esta última expresión era sólo una parte muy menor de la regla de Bayes y debemos coincidir con este autor. Bayes realmente no se preocupó de este problema, sino Laplace. Pearson, dirigiéndose a Burnside (1924), quien se refiere a esta última formulación como "la regla de Bayes", atribuyéndola a Poincaré (1916), escribió: "*Dr. Burnside cites as Bayes' formula, what is only an element in Bayes' Theorem, and he does so on the strength of Poincaré, who in all probability had not studies Bayes´ original work*" y agrega "*Dr. Burnside, who does not seem to have read Bayes*" [1].

Karl Pearson (1857-1956), a quien hemos mencionado repetidas veces, fue un importante matemático británico, padre de la estadística matemática, estudioso de la biometría y famoso editor de la revista Biometrika. Gran parte de los fundamentos de la estadística que usamos hoy se deben a este importante científico. También es famoso por sus frases irónicas y su larga polémica con otro de los grandes pensadores de la estadística: Ronald Fisher. Para Karl Pearson [11] el teorema de Bayes-Laplace, en sus propias palabras, se puede expresar como: "*Bayes' Theorem as we now usually state it is of the following nature. We suppose past experience to be represented by p successes and q failures in n trials, and we ask what is the chance of r successes and s failures in further m trials. We hold this problem answered by the expression*

$$\mathrm{P}\left(B_r^{r+s} \mid B_p^{p+q}\right) = \binom{r+s}{r} \frac{\int_0^1 x^{p+r} (1-x)^{q+s} dx}{\int_0^1 x^p (1-x)^q dx} \tag{14}$$

Como vemos, para Pearson el teorema de Bayes es más bien lo que hoy se conoce como regla de sucesión de Laplace o Bayes-Laplace y no la expresión acotada que usamos hoy.

Ronald A. Fisher (1860-1962), al igual que Pearson, se formó en matemáticas y después derivó a la biología inspirado por artículos de Pearson. Fisher realizó grandes contribuciones a la síntesis del Darwinismo, la teoría evolutiva, la genética de poblaciones y, por supuesto, la estadística. Junto a Pearson hay que considerarlo uno de los padres de la bioestadística. Conocido además por sus largas polémicas con Pearson, es quien dio las bases de la aproximación frecuentista a la docimasia de hipótesis como la utilizamos hoy, sólo levemente modificada por los aportes de Egon Pearson (hijo de Karl) y Jerzy Neyman. Fisher fue un opositor al uso de las probabilidades inversas en inferencia estadística, manifestando que "*admittedly depended upon an arbitrary assumption, so that the whole method has been widely discredited*" y "*inverse probability, which like an impenetrable jungle arrests progress towards precision of statistical concepts*" [12]. Sin embargo, reconoció su uso en la argumentación del método de máximo verosimilitud: "*I must indeed plead guilty in my original statement of the Method of Maximum Likelihood [in 1912] to having based my argument upon the principle of inverse probability; in the same paper, it is true, I emphasized the fact that such inverse probabilities were relative only*". Fisher criticó principalmente el uso de las distribuciones *a priori* uniformes, pero tuvo claridad en que en ocasiones los parámetros pueden ser considerados como variables aleatorias. Fisher es el autor del argumento fiduciario que buscó reemplazar este supuesto y lograr inferencias respecto a parámetros, lo que es hoy la base de nuestros conocidos intervalos de confianza.

Conclusiones

Lo que hoy conocemos como teorema de Bayes solo se encuentra parcialmente en el trabajo original. Laplace tuvo una gran participación en el desarrollo del teorema y de la aplicación a la probabilidad de las causas. El trabajo original de Bayes en realidad fue mucho más allá del teorema que enunciamos hoy, buscando la probabilidad de eventos futuros ante el conocimiento de eventos pasados, lo que propuso una base para las inferencias inductivas. Debemos coincidir con Pearson en la relevancia de la regla o teorema de sucesión de Bayes-Laplace. El constructo teórico de Bayes y Laplace tuvo y tiene repercusiones en el desarrollo de la estadística, tuvo influencia en Fisher en el desarrollo del método de máxima verosimilitud y tiene importancia epistemológica en la forma de probar hipótesis.

Referencias

[1] Dale, A.I. (1982). Bayes or Laplace? An examination of the origin and early applications of Bayes´ theorem. *Archive for History of Exact Sciences*, *27*(1), 23-47. doi.org/10.1007/BF00348352

[2] Poincaré H. (1912). *Calcul des probabilités*. Gauthier-Villars: París.

[3] Stigler, S.M. (1983). Who discovered Bayes´s theorem?. *The American Statistician*, *37*(4): 290-296. doi.org/10.2307/2682766

[4] Bayes, T. Communicated by Mr. Richard Price (1763). An Essay towards solving a problem in the Doctrine of Chances. *Philosophical Transactions*, *53*, 370-418.

[5] Stigler, S.M. (2013). The true title of Bayes´s essay. *Statistical Science*, *28*(3), 283-288. doi.org/10.1214/13-STS438

[6] De Moivre, A. (1718). The doctrine of chances. London.

[7] de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, *7*, 1–68.

[8] Edwards, A.W.F. (1978). Commentary on the arguments of Thomas Bayes. *Scandinavian Journal of Statistics*, *5*(2), 116-118.

[9] Laplace, P.S. (1774). Mémoire sur la probabilité des causes par les événements. *Mémoires de l'Académie royale des Sciences de MI (Savants étrangers)*, *4*: 621-656. Reimpreso en: Laplace, *Oeuvres complètes* (París, Francia: Gauthier-Villars et fils, 1841), *8*, pp. 27-65.

[10] Crépel, P. (1988). Condorcet et léstimation statistique. *Statistique et Analyse des Données*, *13*(2), 43-64.

[11] Pearson, K. *Lectures by Karl Pearson at university College, London, during the academic sessions 1921-1933*. In Pearson, E.S. (ed.) (1978), *The History of Statistics in the 17th and 18th Centuries, against the changing background of intellectual, scientific and religious thought*, Chas. Griffin & Co. Ltd.: London.

[12] Zabell, S. (2022) Fisher, Bayes, and Predictive Inference. *Mathematics*, *10*(10), 1634. doi.org/10.3390/math10101634

Anexo 1

Demostración de la integral definida

$$\int_0^1 \binom{m+n}{n} x^n (1-x)^m dx = \frac{1}{m+n+1}$$

Dejemos a un lado $\binom{m+n}{n} = \frac{(m+n)!}{m!n!}$ y resolvamos por partes.

Sea $u = (1-x)^m$, $du = -m(1-x)^{m-1}dx$. Sea $dv = x^n dx$, $v = \frac{x^{n+1}}{n+1}$. Entonces,

$$\int_0^1 x^n (1-x)^m dx$$

$$= \frac{(1-x)^m x^{n+1}}{n+1}\Bigg|_0^1 + \frac{m}{n+1}\int_0^1 (1-x)^{m-1} x^{n+1} dx$$

$$= \frac{m}{n+1}\int_0^1 (1-x)^{m-1} x^{n+1} dx$$

Repitiendo el proceso,

$$\begin{aligned}&\frac{m}{n+1}\frac{m-1}{n+2}\int_0^1 (1-x)^{m-2}x^{n+2}dx\\ &\quad= \cdots \frac{m(m-1)(m-2)\cdots 1}{(n+1)(n+2)\cdots(n+m)(n+m+1)}\\ &\quad= \frac{m!n!}{(m+n+1)!}\end{aligned}$$

Entonces, la integral propuesta es

$$\frac{(n+m)!}{n!m!}\frac{n!m!}{(n+m+1)!} = \frac{1}{n+m+1} \qquad \square$$

Anexo 2

Repitiendo el mismo procedimiento se puede demostrar que

$$\frac{\int_0^1 x^{n+1}(1-x)^m dx}{\int_0^1 x^n(1-x)^m dx} = \frac{n+1}{m+n+2}$$

y que

$$\begin{aligned}&\frac{\int_0^1 x^{n+r}(1-x)^{m+s}dx}{\int_0^1 x^n(1-x)^m dx}\\ &= \frac{(m+1)(m+2)\cdots(m+s)(n+1)(n+2)\cdots(n+m+1)}{(n+r+1)(n+r+2)\cdots(n+m+r+s+1)}\end{aligned}$$