



1 de noviembre de 2013 | Vol. 14 | Núm. 11 | ISSN 1607 - 6079

# ARTÍCULO

## EL DESARROLLO DE LA ESTADÍSTICA BAYESIANA

*Eduardo Gutiérrez*

## EL DESARROLLO DE LA ESTADÍSTICA BAYESIANA

### Resumen

2013 ha sido denominado el Año Internacional de la Estadística<sup>1</sup> para conmemorar los 300 años de la publicación de *Ars Conjectandi*, obra de Jacob Bernoulli que puede considerarse el origen de la teoría moderna de la probabilidad y la base de la llamada estadística frecuentista. Este año también se conmemoran los 250 años de la publicación de *An essay towards solving a problem in the doctrine of chances*, trabajo atribuido a Thomas Bayes en el que se introduce la noción de probabilidad inversa (en la forma de un resultado conocido hoy como el teorema de Bayes) y que

ha dado origen a la llamada estadística bayesiana. El propósito del presente trabajo es discutir brevemente el desarrollo de este enfoque de inferencia estadística y señalar los aspectos que lo distinguen de los métodos estadísticos tradicionales.

**Palabras clave:** Incertidumbre, Inferencia Estadística, Teorema de Bayes, Teoría de la Decisión, Teoría de la Probabilidad.

### THE DEVELOPMENT OF BAYESIAN STATISTICS

#### Abstract

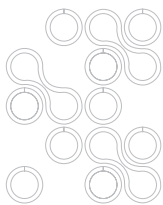
*"2013 has been designated the International Year of Statistics to commemorate the 300th anniversary of the publication of *Ars Conjectandi*, written by Jacob Bernoulli and widely regarded as the origin of modern probability theory and the basis of the so-called frequentist approach to Statistics. This year also commemorates the 250th anniversary of the publication of *An essay towards solving a problem in the doctrine of chances*, a paper attributed to Thomas Bayes which introduced the notion of inverse probability and gave rise to the development of Bayesian Statistics. The purpose of this paper is to briefly describe such development, and to highlight the aspects that distinguish the Bayesian approach from traditional statistical methods..*

**Keywords:** Bayes' Theorem, Decision Theory, Probability Theory, Statistical Inference, Uncertainty.

“

En el caso de la estadística, podemos decir que su objeto de estudio son los fenómenos aleatorios; es decir, aquellos que no se pueden predecir con certeza.

”



[1] <http://www.statistics2013.org/>

## EL DESARROLLO DE LA ESTADÍSTICA BAYESIANA

### Introducción

**A** ¿Qué es la estadística? Quizás la mejor manera de definir una disciplina sea en términos de su objeto de estudio. En el caso de la estadística, podemos decir que su objeto de estudio son los fenómenos aleatorios; es decir, aquellos que no se pueden predecir con certeza. Para poder describir adecuadamente la incertidumbre asociada a este tipo de fenómenos es necesario echar mano del concepto de probabilidad.

En 1654, el Chevalier de Meré le plantea a Blaise Pascal el siguiente problema: ¿cómo dividir las apuestas en un juego de azar que acaba prematuramente si uno de los jugadores lleva la ventaja? Este problema había confundido a los matemáticos durante 200 años. Pascal discutió el problema con Pierre de Fermat, un abogado y matemático aficionado, pero no por ello menos brillante. Su solución al problema sentó las bases de la teoría de la probabilidad. Con el paso de los años, los matemáticos transformaron esta teoría, de una curiosidad de jugadores, en una poderosa herramienta para organizar, interpretar y utilizar información con el fin de tomar mejores decisiones.



En 1713 se publicó un tratado de Jacob Bernoulli en el que introduce la ley de los grandes números. Este resultado puede considerarse la base de la interpretación frecuentista de la probabilidad. Cincuenta años más tarde, Richard Price publicó -de manera póstuma- un trabajo en el que Thomas Bayes introduce el concepto de probabilidad inversa y demuestra una versión del teorema que lleva su nombre. Este trabajo condujo a la interpretación subjetiva de la probabilidad y permitió entender cómo cambian los juicios de una persona acerca de la ocurrencia de un evento a la luz de información (evidencia) adicional.

Consideremos el siguiente ejemplo. Una prueba clínica permite detectar el VIH con una sensibilidad de 95% y una especificidad de 98%. En una población con una prevalencia de VIH de 1/1000, ¿cuál es la probabilidad de que una persona cuya prueba resulta positiva realmente tenga el VIH? Sea  $A$  el evento "la persona tiene el VIH" y  $A^c$  el evento complementario, "la persona no tiene el VIH". Denotemos por  $B$  al evento "la prueba resulta positiva". Entonces, sensibilidad de 95% significa que  $\Pr(B | A) = 0.95$ , mientras que especificidad de 98% quiere decir que  $\Pr(B | A^c) = 0.02$ . Queremos calcular  $\Pr(A | B)$ . El teorema de Bayes nos dice cómo obtenerla:

$$\Pr(A | B) = \Pr(A) \Pr(B | A) / \Pr(B)$$

donde  $\Pr(B) = \Pr(A) \Pr(B | A) + \Pr(A^c) \Pr(B | A^c)$ . En otras palabras

$$\Pr(A | B) = (0.001)(0.95) / 0.021 = 0.045$$

Por lo tanto, ¡más del 95% de las personas cuya prueba resulta positiva en realidad no tienen el VIH! Lo anterior a pesar de las buenas características (sensibilidad y especificidad) de la prueba de diagnóstico.  $\diamond$

A finales del siglo XVIII, Pierre-Simon de Laplace redescubrió el teorema de Bayes y lo aplicó de manera sistemática al análisis de datos. Su trabajo dominó la estadística durante los siguientes 100 años.

## La búsqueda de la objetividad

En el siglo XIX, Karl Pearson estableció los principios de la estadística matemática. Poco después, Francis Galton descubrió el concepto de regresión a la media (que explica por qué mucha gente toma decisiones con la expectativa de que "las cosas volverán a la normalidad"). Galton y Pearson son considerados los padres de la estadística moderna. Debido en parte al trabajo de este último, a finales del siglo XIX se empezó a cuestionar un aspecto fundamental del trabajo de Bayes y de Laplace: la interpretación subjetiva de la probabilidad.

En la década de 1920, Ronald Fisher, por una parte, y Ergon Pearson y Jerzy Neyman, por la otra, propusieron nuevos métodos para evaluar la evidencia contenida en los datos al comparar dos hipótesis distintas. Su enfoque se basa en la interpretación frecuentista de la probabilidad mencionada anteriormente. La idea es indicar con qué frecuencia el investigador puede esperar obtener un resultado determinado si un experimento se repitiera y analizara bajo las mismas condiciones un número grande de veces.

Continuando con nuestro ejemplo, recordemos que  $Pr(B | A) = 0.95$  y  $Pr(B | A^c) = 0.02$ . Si ocurre el evento  $B$ , y dado que  $Pr(B | A) > Pr(B | A^c)$ , se dice que el evento  $A$  es más verosímil que  $A^c$ , si ocurre el evento  $B$ . Es decir, si la prueba resulta positiva, se concluye que es más verosímil que la persona tenga el VIH, frente a la alternativa de que no lo tenga. Este resultado parece contradecir el anterior; sin embargo, aquí el argumento es que, al observar a un número grande de individuos cuyo diagnóstico es conocido, la prueba clínica resultará positiva con mayor frecuencia en aquellos individuos que tienen el VIH que en aquellos individuos que no lo tienen. La conclusión, en este caso, no necesita establecer el valor de  $Pr(A)$ , por lo que se argumenta que este enfoque es "objetivo" (depende exclusivamente del resultado de la prueba y no requiere ninguna referencia externa). Notemos, sin embargo, que desde el punto de vista de la persona que se somete a la prueba es deseable que se tome en cuenta la información contenida en la prevalencia, a través de las probabilidades  $Pr(A)$  y  $Pr(A^c)$ , y que la respuesta sea específica a su situación particular (la probabilidad de tener el VIH dado que su prueba resultó positiva) en lugar de tener simplemente una afirmación genérica sobre la proporción de individuos cuya prueba resultaría positiva en una población de personas infectadas con el VIH.



Tomado de Wikimedia  
 Commons.

Este nuevo enfoque, basado en la noción de verosimilitud y en la interpretación frecuentista de la probabilidad, resultó ser relativamente fácil de aplicar en la práctica. Esto motivó su desarrollo constante durante la mayor parte del siglo XX, así como la aparición de *software* especializado para llevar a cabo los procedimientos más comunes.

## Resurgimiento de la estadística bayesiana

A partir de 1920 también hubo avances que eventualmente permitieron consolidar el enfoque bayesiano de la estadística y producir, entre otras cosas, una teoría formal de inferencia estadística a través de la teoría de la decisión. Concretamente, durante la década de 1920 Frank Ramsey realizó los primeros intentos de axiomatización, mientras que unos años más tarde Bruno de Finetti introdujo el concepto intercambiabilidad, el cual tiene implicaciones muy profundas. Junto con el trabajo de Leonard Savage en la década de 1950, estas ideas permitieron extender el ámbito del principio de la utilidad esperada a situaciones en las que la probabilidad debe interpretarse de manera subjetiva, y de esa manera le dieron una base firme a la teoría bayesiana de la decisión en ambiente de incertidumbre. A finales de esa década, Howard Raiffa y Robert Schlaifer ya estaban aplicando esta teoría a problemas del área de economía, y poco tiempo después Dennis Lindley comenzó a aplicarla sistemáticamente para resolver problemas estadísticos.



Retomando nuestro ejemplo, la aparente contradicción entre el análisis inicial y el basado en la verosimilitud se debe en gran parte a que, en general, nuestra intuición no es suficientemente buena al procesar evidencia probabilística. El punto crucial es ¿de qué manera el resultado de la prueba debe cambiar nuestra opinión sobre el evento de que la persona tenga VIH? La prevalencia de VIH puede pensarse como la probabilidad *a priori* (o inicial) que describe nuestros juicios sobre el evento de que la persona tenga VIH antes de conocer el resultado de la prueba:  $\Pr(A) = 0.001$ . Al observar un resultado positivo, nuestros juicios cambian y la probabilidad del evento se modifica:  $\Pr(A | B) = 0.045$ . Esta es la probabilidad *a posteriori* (o final) que describe nuestra opinión sobre la ocurrencia del evento *A* después de conocer el resultado de la prueba.  $\diamond$

En la actualidad, el uso del teorema de Bayes en pruebas de diagnóstico es bastante común y no causa controversias. En cambio, es mucho más controversial su uso en análisis estadísticos generales, en los que los eventos de interés están en términos de los valores desconocidos de los parámetros de un modelo y por lo tanto se requiere especificar una distribución de probabilidad sobre dichos valores.

De esta forma, mientras los métodos estadísticos tradicionales se preguntan qué nos dicen los datos acerca del valor del parámetro (ignorando con ello toda evidencia externa a los datos), los métodos bayesianos se preguntan explícitamente cómo cambia nuestro estado de información acerca del valor del parámetro (o de cualquier otra cantidad de interés) a la luz de los datos observados.

Una de las principales críticas al enfoque bayesiano es que dos personas cuyas opiniones iniciales difieren podrían llegar a conclusiones distintas a pesar de observar exactamente los mismos datos. Si bien esto es cierto en un sentido estricto, conforme se va acumulando la evidencia, las opiniones de esas dos personas deberán converger. En este sentido, la objetividad a la que aspira la ciencia podría replantearse de manera más realista como un "consenso de subjetividades".

Al admitir información adicional, los métodos bayesianos reconocen que cada problema es distinto y promueven que el procedimiento de análisis se adapte al problema en cuestión. Debido a la popularidad de muchas de las técnicas tradicionales, en la práctica no es raro que se trate de adaptar o modificar el problema original con el fin de que pueda ser analizado usando alguna de esas técnicas. Como consecuencia, los métodos bayesianos tienden a ser más flexibles que sus contrapartes tradicionales.



Tomado de Wikimedia Commons.

La implementación de las técnicas bayesianas usualmente requiere de un esfuerzo computacional muy alto. La mayor parte de este esfuerzo se concentra en el cálculo de ciertas características de la distribución final del parámetro de interés. Por ejemplo, hay que integrar para pasar de una distribución conjunta a una colección de distribuciones marginales que sean útiles para hacer inferencias sobre los parámetros de interés. En la gran mayoría de los problemas las integrales requeridas no pueden resolverse analíticamente, por lo que es necesario contar con métodos numéricos eficientes que permitan calcular o aproximar integrales en varias dimensiones. Afortunadamente, durante la segunda mitad del siglo XX se comenzaron a desarrollar técnicas numéricas flexibles y eficientes basadas en métodos de simulación estocástica. Esto, aunado al desarrollo de la tecnología que ha dado lugar a una mayor capacidad de procesamiento y de almacenamiento de los equipos de cómputo, provocó un auge de los métodos bayesianos del que todavía disfrutan.

La tabla siguiente presenta de manera sucinta los principales aspectos que distinguen a la estadística bayesiana de los métodos estadísticos tradicionales.

Estadística bayesiana	Estadística frecuentista
Interpretación subjetiva de la probabilidad (necesidad de asignar una distribución inicial)	Interpretación frecuentista de la probabilidad
Teoría estadística general, basada en la teoría de la decisión (axiomas de coherencia y cuantificación)	Colección de métodos y técnicas <i>ad hoc</i> para resolver problemas específicos
Coherencia interna	Los procedimientos no son necesariamente consistentes entre sí.
Relativamente difícil de aplicar en la práctica (al menos hasta hace algunos años)	Relativamente fácil de aplicar en la práctica (disponibilidad de <i>software</i> )
Énfasis en métodos numéricos de integración (técnicas sofisticadas de simulación estocástica)	Énfasis en métodos numéricos de optimización y en técnicas tradicionales de Monte Carlo

## Conclusiones

Las siguientes propiedades caracterizan al enfoque bayesiano de la estadística: 1) Uso de información adicional: cada problema es único y tiene su propio contexto, del cual se deriva la información inicial sobre el parámetro (o cualquier otra característica) de interés. 2) Interpretación subjetiva de la probabilidad: se reconoce explícitamente que toda asignación de probabilidades es subjetiva (i.e., dependen del estado de información del individuo que las asigna). No pretende ser un enfoque "objetivo". 3) Coherencia interna: al considerar el valor del parámetro como aleatorio, los métodos bayesianos de inferencia se desarrollan de manera natural a partir de la teoría de la probabilidad y por lo tanto no presentan contradicciones internas.

Como puede verse, la discusión científica y filosófica sobre el uso del teorema de Bayes para hacer inferencia estadística ha durado más de un siglo y se ha centrado en aspectos fundamentales del quehacer científico: cómo analizamos la evidencia, cómo cambia nuestra opinión conforme recibimos más información, y cómo tomamos decisiones en presencia de incertidumbre.



## Bibliografía

- [1] BERNARDO, J.M. & Smith, A.F.M. *Bayesian Theory*. Chichester, Wiley.
- [2] BERNSTEIN, P.L. *Against the Gods: The Remarkable Story of Risk*. New York, Wiley, 1996.
- [3] MALAKOFF, D. "Bayes offers a 'new' way to make sense of numbers" en *Science*, vol. 286, 1999, pp. 1460-1464.
- [4] McGRAYNE, S.B. *The Theory that Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy*. New Haven, Yale University Press, 2011.
- [5] SPIEGELHALTER, D.J., Myles, J.P., Jones, D.R. and Abrams, K.R. "Methods in health service research: An introduction to Bayesian methods in health technology assessment" en *BMJ*, 391, 1999 pp. 508-512.