



Réplica: La larga vida científica que le espera a Thomas Bayes

Rejoinder: The long scientific life that awaits Thomas Bayes

Luis Carlos Silva^a
lcsilva@infomed.sld.cu

En primer lugar, me corresponde agradecer a *Comunicaciones en Estadística* la deferencia de someter mi trabajo en defensa del pensamiento bayesiano al juicio crítico no solo de sus lectores, sino también de tres destacados colegas.

Sus criterios coinciden en respaldar mi convicción de que el razonamiento que conduce al profesor Bunge a descalificar en bloque el enfoque bayesiano resulta inaceptable. Los ejemplos creados por el afamado filósofo y que le autorizarían a tal descalificación, no es que tengan dudosa validez, sino que contienen, simplemente, gruesos errores. Sigo convencido, por tanto, de que no hay que admitir la absurda conclusión de que la probabilidad de padecer sida es igual a la de padecerlo supuesto que se es portador de VIH, ni mucho menos creer en la resurrección, para mantener en alta estima a la inferencia bayesiana. No serán enemigos como Bunge quienes pongan en peligro la larga vida científica que le espera a Thomas Bayes.

Sin embargo, el artículo que motivó estos generosos comentarios va más allá que la simple denuncia de los sofismas puntuales en que incurre Bunge en los ejemplos de su libro “Filosofía para médicos”. Se orientaba también a subrayar otras aseveraciones, acaso menos burdas, pero sumamente controvertibles, en especial aquellas relacionadas con el papel de la subjetividad. Estos desaciertos se inscriben en un estilo de corte doctrinal, presente en buena parte del libro, que los propicia y que en parte los explica. Un juicio más global sobre este problema fue desarrollado en un artículo independiente, publicado en una revista de aliento salubrista y no de estadística (Silva 2013), como corresponde al hecho de que tal estilo no se relacionaba exclusivamente con esta última disciplina.

Poco tengo para discrepar con los tres comentaristas. Me ceñiré por tanto a resaltar, y ocasionalmente complementar brevemente, algunas de sus observaciones en torno a la problemática general en que se inscribe la inferencia bayesiana, alentadas por las ideas de Bunge, pero que entrañan aportes relativamente independientes de ellas. Serán simples pinceladas relacionadas con asuntos que me llamaron la atención.

^aInvestigador titular. Escuela Nacional de Salud Pública. La Habana, Cuba.

La lúcida contribución del profesor Andrés Gutiérrez, enfatiza que “la asignación de las probabilidades previas no puede ser arbitraria” y agrega que “esta tarea debe ser concienzuda y responsable”. Desde luego, comparto el espíritu de esta opinión; pero me gustaría insertar una matización. Desde una perspectiva estricta, yo diría que sí “puede” ser arbitraria (siempre que cumpla con los axiomas de Kolmogorov). Lo que ocurre, y creo que es lo que en esencia quiere señalar mi colega, es que en tal caso se corre un riesgo nada desdeñable de que dicha asignación no sea fructífera. Es por ello que ha de ser concienzuda. Asumo el riesgo de parecer descontextualizado y pongo un ejemplo, en apariencia totalmente ajeno a las probabilidades, pero hartamente elocuente.

A mediados de la década de los sesenta del siglo pasado surgió un atleta del salto alto que empleaba un estilo totalmente heterodoxo. Dick Fosbury no acometía el listón corriendo de manera oblicua a él y colocando el cuerpo boca abajo tras la carrera, tal y como hacían todos. Así había conseguido, por ejemplo, el soviético Valery Brumel la marca mundial entonces vigente de 228 cm. Fosbury corría hacia el listón siguiendo una trayectoria curva; una vez frente a la varilla, giraba y se elevaba de espaldas a ella con el brazo más próximo extendido. Con esa técnica Fosbury obtuvo una plaza para representar a Estados Unidos en los Juegos Olímpicos de México, en 1968. Antes del certamen, su método fue objeto de críticas y hasta de burlas, por ser considerado esnobista, rocambolésco y poco natural. Sin embargo, no solo ganó la medalla dorada en México con la mejor marca de la historia de dichos juegos (225 cm) sino que su procedimiento, desde entonces conocido como el “estilo Fosbury”, ha sido el empleado por todos los saltadores posteriores. Resulta expresivo que Dick Fosbury jamás pudo superar la marca de Brumel pero, usando su invención, esta no demoró en caer una y otra vez hasta llegar a la plusmarca mundial vigente, que ostenta el cubano Javier Sotomayor (245 cm)¹. Es decir, el éxito del norteamericano no se debió a que fuera un atleta superior al resto, sino a que usaba un método más eficiente.

La moraleja es muy clara: a la hora de valorar un método, ninguna consideración teórica, ninguna tradición, ningún presupuesto filosófico puede ser más persuasivo que las corroboraciones prácticas de que por su conducto se consigue aquello que se quiere alcanzar. Llevado a nuestro debate, si el manejo de la teoría bayesiana –empleada de una manera sensata y racional– produce resultados palmarios, como efectivamente ocurre, queda poco margen para las objeciones doctrinarias. El profesor Gutiérrez introduce e ilustra la violación del “principio de verosimilitud” en la que se incurren las técnicas inferenciales clásicas. Expresado de manera laxa (para profundizar en el tema, véase Royall (1997)) dicho principio establece que las inferencias estadísticas deberían realizarse a partir y solo a partir de los datos observados. Siendo, como es, una regla difícilmente objetable, siempre me ha resultado fascinante que hayamos convivido durante casi un siglo con tal violación, presente en el núcleo de los ubicuos “valores p ”. Quizás la exposición más fácilmente comprensible de esta inconsecuencia es la debida Berger & Berry (1988). Gutiérrez adiciona que: «El diseño en la recolección de los datos es in-

¹Al final, permítaseme este guiño, los saltadores optaron por el estilo Fosbury no porque les garantizaría el éxito, sino porque aumentaría la probabilidad de tenerlo.

formativo (y no ignorable), no solo en diseño experimental, sino en el análisis de encuestas probabilísticas, para las cuales las propiedades de los estimadores están supeditadas a la medida de probabilidad discreta inducida por el diseño muestral». Dicho en términos algo más simples, se resalta un hecho real: cuando se realiza una inferencia, por ejemplo al calcular un intervalo de confianza a partir de datos obtenidos a través de una muestra probabilística, las fórmulas correctas no solo han de contemplar los datos propiamente dichos, sino también el diseño muestral empleado. Sin embargo, estimo que hay un matiz que distingue ambas situaciones y que merece ser tenido en cuenta: en el cálculo de un “valor p ” interviene, por definición, el diseño del estudio, pues dicho cálculo exige tener en cuenta lo que “podría” haberse observado y no solo lo que objetivamente se observó. Siendo así, dos analistas podrían llegar a diferentes valores p en virtud del plan según el cual cada uno se condujo, no obstante haber obtenido los mismos resultados con el mismo *modus operandi*. Sin embargo, para el cómputo del intervalo de confianza con datos procedentes de una muestra probabilística, si bien se tiene en cuenta algo más que los meros datos, ese algo no depende de lo que el investigador tenía en mente realizar sino de lo que objetivamente hizo. Dos analistas diferentes que conozcan lo que se hizo (el diseño muestral aplicado, en este caso) llegarán al mismo resultado².

Aparte de resaltar el vuelo de su discurso, vertebrado en torno a muy actualizadas referencias bibliográficas, destaco la idea final de Gutiérrez («desconocer el desarrollo científico que la estadística bayesiana ha traído consigo, ya no es una opción»), tan lapidaria como atinada.

El profesor Jairo Fúquene ofrece un ejemplo complementario para desembocar en la constatación, acaso más persuasiva que cualquier disquisición teórica, del impacto que ha tenido el uso de la estadística bayesiana en diversos terrenos (bioestadística, medicina, bioinformática, neurología y genética, para citar solo las disciplinas que él menciona). Me permito volver a sugerir en este contexto la lectura de un libro excepcional, que ya mencioné en mi contribución inicial, pero que en el ínterin ha sido traducido y publicado en castellano con el título *La teoría que nunca murió* (McGrayne 2012).

«No siempre el respaldo de datos como fuente de información previa es sinónimo de objetividad» nos dice el profesor Jorge Ortiz. Excelente acotación. Y luego sintetiza elegantemente la prédica bungeana cuando recuerda que el filósofo «ataca decididamente la subjetividad y arremete furiosamente contra los bayesianos cuando observa que le abren la puerta?». Bunge aprecia el impetuoso desarrollo del pensamiento bayesiano, señala Ortiz, «como una amenaza al progreso de la ciencia y especialmente de la medicina».

La objetividad es una aspiración legítima del pensamiento científico, pero conse-

²El asunto es sutil, y no es fácil de exponer en pocas líneas. En otro sitio (Silva 2009) desarrolló un sencillo ejemplo que muestra que una persona que esté al tanto de lo que un investigador hizo –desde el comienzo hasta el final de su experimento– así como de los datos obtenidos, no sabría cómo calcular el valor p hasta que dicho investigador no le comunique cuál era el plan experimental al que respondía eso que hizo, el cual solo se halla en su cabeza.

guirla en estado puro es una quimera. Contribuir a alcanzarla es (casi) la razón de ser de la estadística. Pero garantizarla por su conducto es imposible. Consecuentemente, desde una óptica fundamentalista, toda la estadística inferencial, y no solo la bayesiana, debería ser erradicada, ya que lo que pudiéramos llamar la estadística clásica” está plagada de subjetivismo. El ejemplo más inmediato es el umbral empleado en las pruebas de hipótesis para declarar “significación”: el ubicuo $\alpha = 0.05$ es intrínsecamente subjetivo. Las propias pruebas de significación utilizadas pueden ser paramétricas o no paramétricas, de una o de dos colas, desarrollarse con o sin correcciones (de Yates o de Bonferroni, por ejemplo), sin que exista un criterio indiscutible para elegir unas u otras. La elección del tamaño muestral adecuado para un estudio inexorablemente demanda del concurso de la subjetividad; una detallada demostración del carácter hondamente subjetivo de tal determinación puede hallarse en Silva (2000) o en Schulz & Grimes (2005). Igualmente subjetiva es la decisión de cuáles son las variables iniciales de un modelo de regresión múltiple antes de aplicar una selección “paso a paso”, o la de aplicar dicha selección hacia adelante o hacia atrás. Otro tanto ocurre con la valoración de si cierto valor se debe o no considerar “aberrante” (un *outlier*), o con la decisión de cuáles trabajos han de incluirse en un metanálisis. La lista podría continuarse, pero se trata de una verdad tan obvia que no vale la pena extenderse. No casualmente, Berger & Berry (1988) en el contexto de la contribución ya citada, escribían hace 25 años: «el uso común de la estadística parece haberse fosilizado, principalmente debido a la visión de que la estadística clásica es la forma objetiva de analizar datos».

El profesor Ortiz resume su intervención diciendo que «En medio del ataque de Bunge, percibo una defensa de mayor dimensión: proteger a la estadística, bayesiana o no, del uso irracional o intencionado de la subjetividad». Puede ser que tan loable propósito haya estado en el ánimo de Bunge. Me parece generosa la buena voluntad del profesor Ortiz al hacer esa lectura; pero me resulta difícil admitir que se quiera proteger a la estadística bayesiana del uso irracional de la subjetividad mediante el recurso de embestir contra ella usando argumentos irracionales.

Recibido: 1 de septiembre de 2013

Aceptado: 7 de octubre de 2013

Referencias

- Berger, J. O. & Berry, D. A. (1988), ‘Statistical analysis and the illusion of objectivity’, *The American Scientist* **76**, 159–165.
- McGrayne, S. B. (2012), *La teoría que nunca murió: de cómo la Regla de Bayes permitió descifrar el código Enigma, perseguir los submarinos rusos y emerger triunfante de dos siglos de controversia*, Grupo Planeta.
- Royall, R. M. (1997), *Statistical evidence: a likelihood paradigm*, Chapman & Hall/CRC, Boca Ratón.

- Schulz, K. F. & Grimes, D. A. (2005), 'Sample size calculations in randomised trials: mandatory and mystical', *The Lancet* **365**, 1348–1353.
- Silva, L. C. (2000), *Diseño razonado de muestras y captación de datos para la investigación sanitaria*, Díaz de Santos, Madrid.
- Silva, L. C. (2009), *Los laberintos de la investigación biomédica. En defensa de la racionalidad para la ciencia del Siglo XXI*, Díaz de Santos, Madrid.
- Silva, L. C. (2013), 'Reflexiones a raíz de *Filosofía para médicos*, un texto de Mario Bunge', *Salud Colectiva* **9**(1), 115–128.