

LA ALTERNATIVA BAYESIANA

Luis Carlos Silva Ayçaguer
Instituto Superior de Ciencias Médicas. La Habana.

Quizás algunos lectores de este boletín han oído hablar del *posmodernismo* y del *deconstructivismo*. Es posible que les ocurra como a mí, que solo tengan una idea vaga de lo que significan estos términos. Parece tratarse de una corriente que sostiene la necesidad de “desmontar” (que no destruir) viejos paradigmas artísticos, sociales o políticos, y suplirlos por representaciones más realistas.

La inferencia estadística clásica, en particular la teoría de pruebas de hipótesis, no entra al nuevo siglo en su mejor momento. Las observaciones críticas que se han venido acumulando desde su creación conforman hoy un reclamo metodológico alternativo de tal magnitud que cada día se torna más difícil de soslayar. Me permitiré hacer un pequeño ejercicio de “deconstrucción” al respecto, y de inmediato esbozar un enfoque alternativo para el análisis estadístico de datos que, en cierto sentido, se contrapone al enfoque de la “estadística frecuentista”.

Para desarrollar estas ideas, coloquémonos en el escenario más clásico: un investigador realiza un ensayo clínico para valorar si cierto tratamiento novedoso para enfrentar una enfermedad dada produce una tasa de recuperación más alta que el tratamiento convencional. Planifica valorar sus resultados mediante una prueba de hipótesis convencional. Pero, ¿qué problemas le esperan? Sintetizaré las asechanzas en cinco puntos.

1. Si la verdadera diferencia que se está evaluando no es nula (que es lo que debe ocurrir, pues si los tratamientos son diferentes no hay razón alguna para esperar que tengan efectos EXACTAMENTE iguales), por muy pequeña que sea esa diferencia, siempre se podrá rechazar la hipótesis de nulidad si se trabaja con una muestra suficientemente grande. Esta es una muy grave imputación, pues lo que nos dice es que la decisión queda en manos de un elemento exógeno a la realidad que se examina (la respuesta a la pregunta formulada depende más de los recursos de que dispongamos que de la realidad estudiada). Dicho de otro modo: si la muestra es muy grande, las pruebas de significación son inútiles, pues ya se sabe el resultado que arrojarán.
2. Las pruebas de significación constituyen supuestamente un proceso de decisión sobre una hipótesis (que se ha de aceptar o no); pero el pensamiento científico no discurre así realmente. No podemos ver una hipótesis como si fuera una invitación al cine, que aceptamos o rechazamos. Nuestras convicciones científicas pueden ser más o ser menos firmes, pero siempre son provisionales, y nuestras representaciones de la realidad tienen en cada momento un cierto grado de credibilidad abierto a cambios y perfeccionamientos en la medida que nuevos datos lo aconsejen. Lo ideal sería contar con un procedimiento que, en lugar de desechar o no una hipótesis, sirva para “poner al día” la opinión que ella nos merece a la luz de dichos datos.
3. La información anterior a los datos, proveniente de estudios previos o de la experiencia empírica informalmente acumulada (que siempre se tiene sobre el problema que se examina) no interviene en el modelo estadístico de las pruebas de hipótesis. Este supone un vacío de opiniones y percepciones que no condice con la práctica real.
4. Si el investigador obtiene, por ejemplo, que el porcentaje de recuperación para la terapia convencional es del 45% y que para el tratamiento experimental asciende a 66%, pero la prueba que emplea arroja, pongamos por caso, una $p=0,12$, típicamente el investigador no tiene más remedio que comunicar que “no hay significación”, que “no hay suficiente evidencia muestral como para afirmar que los dos tratamientos difieren entre sí”. Sin embargo, es posible que agregue (con todo derecho,

por cierto): "pero obsérvese que la diferencia es notable; sí hubiéramos trabajado con una muestra mayor, muy posiblemente sí hubiéramos encontrado significación" El investigador se resiste a desdeñar totalmente el resultado (que es lo que en rigor tendría que hacer de acuerdo con las reglas del método que emplea), se rebela ante la idea de actuar como si esa diferencia del 21% estuviera diciendo lo mismo que una diferencia de, por ejemplo, 2%.

5. Puesto que la p que se calcula resume la probabilidad, supuesta válida la hipótesis nula, de haber observado lo que efectivamente se observó más lo que se podría haber observado en la dirección contraria a la hipótesis nula, el valor de p no solo depende de los datos y de la hipótesis que se valora, sino también de cómo se haya concebido el experimento. Dos diseños diferentes pueden dar lugar al mismo resultado empírico pero a distintos valores de p . Este es un detalle técnico, acaso demasiado sutil salvo para estadísticos profesionales; pero no por ello menos elocuente.

Es natural que se aspire a contar con un procedimiento inferencial que tenga los siguientes rasgos (ninguno de los cuales, como se ha dicho, está presente en la teoría frecuentista de las pruebas de hipótesis):

- a. Que cuanto mayor sea el tamaño muestral, con más elementos se cuente para valorar adecuadamente la realidad que esa muestra representa.
- b. Que valore la credibilidad o verosimilitud de las hipótesis en lugar de obligarnos a adoptar decisiones dicotómicas sobre ellas.
- c. Que no parta de una supuesta orfandad total de información; que el modelo de análisis contemple de manera formal el conocimiento previo y la experiencia precedente
- d. Que no desdeñe nunca el resultado de un estudio. Un método que otorgue a los datos la importancia que tienen, ni más ni menos, pero que en ningún caso conduzca a actuar como si no se hubiera hecho el experimento o la observación. Además, que no se rija por umbrales mágicos por encima de los cuales un resultado es demostrativo de algo, y por debajo del cual no nos dice nada.
- e. Que el resultado que se calcule a partir de los datos para sacar conclusiones no dependa de una información que no está presente en los datos propiamente dichos tal cómo la forma en que se haya diseñado el experimento

Pues bien, el enfoque bayesiano se precia de cumplir con las cinco condiciones. Tanto en el ámbito epidemiológico como en el clínico es bien conocido el teorema de Bayes por su utilidad para la valoración de pruebas diagnósticas. Pero el interés por el enfoque bayesiano va mucho más allá que esta conocida aplicación. Al admitir un manejo subjetivo de la probabilidad, el analista bayesiano queda habilitado para emitir juicios de probabilidad sobre la validez de una hipótesis H y expresar por esa vía su grado de convicción al respecto. En su versión más elemental y en este contexto, el Teorema de Bayes asume la forma siguiente:

$$P(H | \text{datos}) = \frac{P(\text{datos} | H)}{P(\text{datos})} P(H)$$

La probabilidad *a priori* de una hipótesis, $P(H)$, es la que se fija o sugiere antes de tener en cuenta los datos (por eso es *a priori*). Una vez realizada una observación o un experimento, se incorpora la evidencia que aportan los datos al análisis a través del teorema y se obtiene una probabilidad *a posteriori*, $P(H|\text{datos})$. $P(H)$ y $P(H|\text{datos})$ representan en principio números; sin embargo, también $P(H)$ pudiera representar nuestra convicción inicial expresada a través de una distribución de probabilidades. En tal caso, una vez observados los datos, el teorema "devuelve" también una distribución, que ha de interpretarse como la nueva percepción, ahora "actualizada" por los datos.

Los métodos bayesianos han sido impugnados con el argumento de que, al involucrar la visión individual que el investigador tiene *a priori*, se introduce un componente de subjetividad en el

análisis. Cabe advertir, por una parte, que el enfoque frecuentista no está exento de decisiones subjetivas (el propio nivel de significación es el ejemplo más obvio); pero por otra parte, y esto es mucho más importante, la subjetividad (entendida como algo bien diferente de la arbitrariedad o el capricho) es un fenómeno inevitable, especialmente en el obligado marco de incertidumbre en que operan las ciencias médicas, biológicas, psicológicas y sociales. Por otra parte, ha de subrayarse que el enfoque bayesiano es en ese sentido más honesto y transparente, pues no solo admite la subjetividad como un hecho inevitable sino que comunica de modo explícito la forma que esta asume.

Aunque las bases de la estadística bayesiana datan de hace más de dos siglos, es ahora que empieza a asistirse a un uso apreciable de este enfoque en la investigación biomédica. Una de las razones que explican tal realidad y que a la vez augura un prominente futuro, es que algunos de los problemas de cierta complejidad exigen posibilidades computacionales accesibles solo ahora para el común de los investigadores. Acaso solo la era informática es capaz de consentir “desconstrucciones” capaces de dar a luz nuevos paradigmas para la próxima centuria.