



Enfermería Investiga

Investigación, Vinculación, Docencia y Gestión

<https://revistas.uta.edu.ec/erevista/index.php/enfi>



EL ENFOQUE BAYESIANO UNA RENOVADA FORMA DE INFERIR THE BAYESIAN APPROACH A RENEWED WAY OF INFERRING

Jesús Estévez¹ <https://orcid.org/0000-0001-9251-8488>

¹Profesor Titular. Instituto de Investigaciones Clínicas, Facultad de Medicina, Universidad del Zulia, Maracaibo/Venezuela.

2477-9172 / 2550-6692 Derechos Reservados © 2022 Universidad Técnica de Ambato, Carrera de Enfermería. Este es un artículo de acceso abierto distribuido bajo los términos de la Licencia Creative Commons, que permite uso ilimitado, distribución y reproducción en cualquier medio, siempre que la obra original es debidamente citada.

Recibido: 03 de octubre 2022

Aceptado: 21 de noviembre 2022

Los procesos estadísticos entendidos como la recolección y clasificación de datos tienen una historia muy antigua. Muchos pueblos en los albores de la civilización implementaron recolección de datos de asuntos muy diversos, censos de población, inventarios de recursos, dotaciones militares, etc., con la finalidad de planificación, control y toma de decisiones sobre las actividades y recursos de esas sociedades.

Esos esbozos de la disciplina estadística carecían de herramientas matemáticas sofisticadas. Estas herramientas como la algebra al inicio hasta el cálculo integral nacieron al unísono de estos primeros ensayos estadísticos, pero sin conexión inicial con los mismos. Pasarían siglos antes de que estas herramientas matemáticas se integraran a los procesos de recolección de datos, su integración y desarrollo en este ámbito llevó a la creación del método científico y la estructuración de la ciencia de análisis de datos e inferencia moderna (1). Actualmente estos procesos como muestreo, análisis de datos e inferencia son fundamentales en la búsqueda de la realidad de los fenómenos estudiados.

Para implementar y hacer eficientes tales procesos emergieron una variedad de visiones sobre cómo abordarlos. La principal visión que ha reinado por décadas en la ciencia ha sido la llamada estadística frecuentista o clásica, definida sobre la distribución de frecuencias, el teorema del valor central y la probabilidad de la hipótesis nula, es una estadística considerada cerrada, pues solo utiliza para sus procesos analíticos la información proveniente de los datos recolectados. Aunque todo proyecto de investigación requiere hacer una revisión bibliográfica de las investigaciones previas del fenómeno a estudiar, los hallazgos de esos estudios no forman parte del análisis e inferencia en la estadística frecuentista (2). El método de esta estadística ha sido mejorado para paliar ciertas debilidades del modelo, pero los paradigmas principales siguen incólumes (3), esta solidez o rigidez puede ser interpretada como un estancamiento, al perder el carácter perfectible que debe tener toda ciencia.

Dentro de los desarrollos estadísticos existen otras maneras de abordar el problema de la probabilidad y de la significancia de las hipótesis. Entre ellas, la surgida del Teorema de Bayes, la estadística bayesiana, que en términos prácticos implica que la probabilidad de un fenómeno determinado es condicionada por

el conocimiento y las creencias que se tienen sobre el referido fenómeno. Este teorema describe en términos matemáticos el proceso usual de adquisición de conocimiento, normalmente no se afronta el estudio de un fenómeno sin una base conceptual del mismo, sea esta teórica, empírica o fáctica. Partir de cero, contraviene el esquema natural donde a las creencias o conocimientos previos sobre un determinado fenómeno, se suman las nuevas experiencias, disminuyendo su interacción con las primeras, el nivel de incertidumbre entre el concepto del fenómeno y la realidad de este (4,5).

La lógica del pensamiento bayesiano se puede resumir en un ejemplo hipotético muy simple: Se ejercita al alba diariamente en un mismo trayecto, pasando frente a un edificio de apartamentos, que colinda con una casa cuyo techo está a la par del balcón del 2do piso del edificio referido. Siempre que se pasa por el frente, la puerta y las ventanas de ese balcón están cerradas. Por otra parte, los habitantes de la casa siempre están activos a esa hora y hay estacionados 2 taxis (esa sería la experiencia previa). Un día al pasar se observa que la puerta del balcón está abierta, no hay presente ninguna persona y hay oscuridad en su interior, extrañamente, los vecinos no se ven por ninguna parte y los taxis no están. Esta sería la nueva experiencia. ¿Qué inferencia puede hacerse de esta situación y sobre qué bases? ¿Con qué porcentaje de certeza podría inferirse que se ha cometido un robo, que los autores son los vecinos y que los taxis han sido usados para llevarse el botín del robo? Si se infiere de los datos del momento, sin adicionar la experiencia previa, el porcentaje de certeza será muy bajo. Mientras que al adicionar dentro del análisis la experiencia previa, el porcentaje de certeza aumenta exponencialmente. He allí la clave del pensamiento bayesiano, donde la probabilidad de la experiencia previa modifica la probabilidad de la experiencia nueva, siendo a su vez la primera modificada con la adición de conocimiento sobre el fenómeno observado, permitiendo una inferencia más precisa y con menor nivel de incertidumbre (4).

A diferencia de la estadística frecuentista que no asimila ninguna experiencia ajena al experimento, la estadística bayesiana utiliza un elemento subjetivo externo al marco experimental, la experiencia o conocimiento previo sobre el fenómeno (4,6). Esa información previa es la base de la generación de una distribución a priori, donde se describe el comportamiento

probable de la variable acorde al cuerpo de conocimientos, experticias o creencias que se tenga hasta ese momento. La distribución a priori sirve como una plantilla de referencia, en donde se contrasta la distribución de los valores observados, de la relación y análisis de ambas distribuciones se inferirá el incremento o disminución de la probabilidad o certidumbre del marco de las hipótesis planteadas. Las distribuciones a priori pueden ser no informativas donde no hay estudios previos o una opinión especializada del tema o informativas donde si hay conocimientos previos, experiencias empíricas u opinión especializada (5,6).

El manejo de las probabilidades de las hipótesis marca una diferencia entre ambas visiones. La frecuentista calcula solo la probabilidad de que la hipótesis nula sea cierta o no, a través de la prueba de significancia estadística de la hipótesis nula. Del valor de la "p" obtenida, se derivará la aceptación o el rechazo de la hipótesis nula, debe recordarse, pues es un frecuente error de sobrestimación de la "p", que el rechazo de la hipótesis nula no representa ningún apoyo para la hipótesis alternativa (7). La "p" solo es una posibilidad, para conocer su precisión se necesita el intervalo de confianza que indica la amplitud del rango de los parámetros calculados, donde probablemente se encuentre el valor real y confirmando la significancia o no del resultado. Cada punto del intervalo de confianza tiene la misma posibilidad de ser el valor real buscado. La inferencia frecuentista es muy rígida y solo da respuestas dicotómicas, si o no, sin nada intermedio (5,7).

La bayesiana aborda el cálculo de la probabilidad de ambas hipótesis, la nula y la alternativa (pueden ser más de 2 hipótesis). Con el cálculo de ambas probabilidades se tiene una idea del peso específico de cada hipótesis respecto a describir el fenómeno estudiado con mayor precisión. Con los datos observados se calcula la probabilidad de compatibilidad tanto de la hipótesis nula como de la alternativa. El cociente de esta probabilidad de la hipótesis alternativa sobre la hipótesis nula es lo que se conoce como Factor Bayes, cuyo valor determinará que hipótesis recibe mayor apoyo del estudio, donde un valor de 1 indicaría no diferencia, un valor >1 apoya mayormente a la hipótesis alternativa, mientras que un valor <1 la apoyada será la hipótesis nula. Se calcula además un índice de creencia o credibilidad que contiene el más probable valor del parámetro estudiado y un rango de dispersión. Por último, se calcula la probabilidad a posteriori o actualizada, donde se determina la modificación de la probabilidad a priori producida por el conocimiento adicional (5,8).

La inferencia bayesiana es flexible pues no se aferra al valor de la "p", no responde dicotómicamente sino de forma holística. Por ejemplo, la probabilidad de un paciente, de una determinada población, de estar infectado por un X virus es del 0,004. Luego de obtener un examen positivo para el virus, esa probabilidad calculada por métodos bayesianos sube al 0,12. Para el frecuentismo ese pequeño aumento no tiene gran significado, mientras que para el bayesiano indica un dato importante, pues

la probabilidad se ha incrementado 24 veces y supone una pequeña probabilidad de estar infectado con una muy alta probabilidad de un examen positivo cuando se tiene la infección. Un segundo examen positivo, actualiza la probabilidad a 0,798, que es una fuerte evidencia a favor de la infección y corrobora la inferencia previa (9). La versatilidad del método bayesiano logra deducciones aun ante pequeñas magnitudes.

La definición de subjetividad también diferencia ambos enfoques. La frecuentista considera que la objetividad está solo en los datos y que todo elemento añadido fuera de estos es una subjetividad que alterará el análisis (10). El enfoque bayesiano, por el contrario, considera la subjetividad de las experiencias previas o creencias como parte natural del proceso de adquisición de conocimiento y por ende su inclusión es vital en el análisis. La objetividad es un ideal, que se expresa en las variaciones de la probabilidad al actualizarse (a posteriori) tras cada nuevo ensayo (4,6). Al final, la discusión sobre subjetividad es una falacia, pues la base de los cálculos estadísticos está llena de acuerdos como el "α" y todos esos acuerdos son subjetivos (11)

La polémica surgida sobre cuál de estos dos sistemas es mejor, resulta algo estéril, ya que son dos herramientas que bien pueden ser utilizadas alternativamente, en función de que el estudio a ser desarrollado se adecue mejor a uno de los sistemas. Mientras la estadística frecuentista es idónea para grandes muestras, muy restringidos en el control de variables y sin ningún conocimiento previo. La estadística bayesiana es adecuada para estudios con muestras pequeñas, con conocimientos previos, experticias empíricas u opinión de expertos (10,11).

La estadística frecuentista se ha utilizado ampliamente debido a su procedimiento simple y a sus cálculos menos complejos, la mayoría de los programas estadísticos fueron estructurados en función de esa estadística. Mientras que la estadística bayesiana, debido a lo complejo de sus cálculos, hasta hace pocas décadas estaba solo al alcance de matemáticos y estadísticos. La aparición de programas nuevos y antiguos con módulos especiales para la estadística bayesiana, como el paquete MCMC del lenguaje R, SPSS, OpenBugs, JASP, WinBugs, Minitab, han permitido el acceso a esta, de investigadores sin amplios conocimientos matemáticos, como es el caso en el área de las ciencias de la salud (9,12).

Esta es una oportunidad para experimentar con un diferente enfoque de inferir, y liberarnos del anquilosamiento que una sola visión produce en todo desarrollo humano. Eso no significa que desechamos la metodología previa, sino que la complementamos, ampliando los horizontes de la inferencia con diversidad de herramientas que permiten aclarar el universo fenomenológico que nos rodea. No hay q temer avanzar, hay que temer estancarse.

REFERENCIAS BIBLIOGRÁFICAS

1. Feinberg SE. A brief history of statistics in three and one-half chapters: A review essay. *Statistical Science*, 1992; 7(2): 208-225. DOI: 10.1214/ss/1177011360
2. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol*, 2016; 31(4):337-350. DOI: 10.1007/s10654-016-0149-3
3. Nurminen M. Statistical significance-a misconstrued notion in medical research. *Scand J Work Environ Health*, 1997; 23(3):232-235. DOI: 10.5271/sjweh.204
4. Cohen AL, Sidlowski S, Staub A. Beliefs and Bayesian reasoning. *Psychon Bull Rev*, 2017; 24(3):972-978. DOI: 10.3758/s13423-016-1161-z
5. Correa Morales JC, Barrera Causil CJ. Introducción a la estadística bayesiana. Editorial Fondo Editorial ITM, 2018, pp 222. DOI: 10.22430/9789585414242
6. Dienes Z, Mclatchie N. Four reasons to prefer Bayesian analyses over significance testing. *Psychon Bull Rev*, 2018; 25(1):207-218. DOI: 10.3758/s13423-017-1266-z
7. Goodman SN. Toward evidence-based medical statistics 1: The p value fallacy. *Ann Intern Med*, 1999; 130(12): 995-1004. DOI: 10.7326/0003-4819-130-12-199906150-00008
8. Steven N. Goodman SN. Toward evidence-based medical statistics. 2: The Bayes Factor. *Ann Intern Med*, 1999; 130(12):1005-1013. DOI: 10.7326/0003-4819-130-12-199906150-00019
9. Armero C., Rodríguez P., de la Torre Hernández JM. Una pequeña mirada a la estadística bayesiana en el análisis de datos cardiológicos. *REC Interv Cardiol*, 2022; 4(3): 207-2015. DOI: 10.24875/RECIC.M22000284
10. Teira D. Frequentist versus bayesian clinical trials in Philosophy of Medicine V 16 in Handbook of the Philosophy of Science Edit Fred Gifford, 2011, pp 255-297. DOI: 10.1016/B978-0-444-51787-6.50010-6
11. Silva LC, Muñoz A. Debate sobre métodos frecuentistas vs bayesianos. *Gac Sanit*, 2000; 14(6): 482-494. DOI: 10.1016/S0213-9111(00)71916-9
12. Lilford RJ, Braunholtz D. Who's afraid of Thomas Bayes? *J Epidemiol Comm Health*, 2000; 54(10): 731-739. DOI: 10.1136/jech.54.10.731