



Investigación en Ciencias Sociales y Humanas: ¿Frecuentista O Bayesiana?

Research in Social and Human Sciences: Frequentist Or Bayesian?

Fabrizio O. Penna

fabrizio.penna@gmail.com

Especialista en Investigación en Ciencias Sociales y Humanas. Profesor Asociado de Investigación Educativa I - Lic. y Prof. en Cs. de la Educación - Facultad de Ciencias Humanas e Investigador (PROICO 12-0220 "Estrés y salud: Una Propuesta Integradora para la Promoción, Prevención e Intervención de Problemas Psicosociales, Orientado a un Desarrollo Saludable de la Persona y la Comunidad" –FaPsi). Universidad Nacional de San Luis.

70

O. Hernán Cobos

ohcobos@gmail.com

Especialista en Metodología de la Investigación Científica. Profesor Adjunto de Investigación Educativa I - Lic. y Prof. en Cs. de la Educación - Facultad de Ciencias Humanas e Investigador (PROICO 12-0220 "Estrés y salud: Una Propuesta Integradora para la Promoción, Prevención e Intervención de Problemas Psicosociales, Orientado a un Desarrollo Saludable de la Persona y la Comunidad" – FaPsi). Universidad Nacional de San Luis.

Sebastián M. Vázquez Ferrero

lasagnad@gmail.com



Dr. en Psicología. Jefe de Trabajos Prácticos de Investigación Educativa I-Lic. y Prof. en Cs. de la Educación - Facultad de Ciencias Humanas e Investigador (PROICO 12-0718 "La investigación en psicología y su incidencia en la formación del psicólogo" – FaPsi). Universidad Nacional de San Luis.

Adriano Penna

apenna99@gmail.com

Estudiante de la carrera Licenciatura en Análisis y Gestión de Datos. Facultad de Ciencias Físico Matemáticas y Naturales – Facultad de Ciencias Económicas, Jurídicas y Sociales. Universidad Nacional de San Luis.

Resumen

El presente trabajo no pretende ser una solución a la pregunta planteada en el título sobre cuál es la mejor -o la más conveniente- de las miradas o enfoques, ya sea frecuentista o bayesiana, cuando se investiga (desde un enfoque pospositivista) en las Ciencias Sociales y Humanas, sino que procura un acercamiento, lo más descriptivo posible, de ambas teorías.

Palabras claves: enfoque frecuentista, enfoque bayesiano, Ciencias Sociales y Humanas

Abstract

This paper does not pretend to be a solution to the question posed in the title about which is the best - or the most convenient - of the approaches, either frequentist or Bayesian, when researching (from a post-positivist approach) in the Social and Human Sciences, but it tries to provide an approach, as descriptive as possible, of both theories.



Keywords: frequentist approach, Bayesian approach, Social and Human Sciences

Introducción

Como bien recordarán, la estadística -per se- presenta dos grandes aristas: descriptiva e inferencial. Y, parafraseando a Russell (1989), podemos pensar a la probabilidad como la pubertad que conecta a la infancia de la estadística descriptiva con la madurez de la estadística inferencial.

Es claro que, cuando pensamos en investigación en Ciencias Sociales y Humanas, en nuestras instituciones académicas se da, mayoritariamente, un enfoque frecuentista. Donde, y en caso de darse, se ve muy por encima (lo cual hace que -casi- se pierda) todo lo referido al enfoque bayesiano.

Para comenzar, podemos señalar que ambas probabilidades, tanto la frecuentista como la bayesiana, son dos enfoques diferentes para calcular y entender la probabilidad en el contexto de la teoría del cálculo de probabilidades. Por ello, y no está de más decirlo, de acuerdo a lo expresado por Ruiz-Benito et al. (2018, p.136), que “la teoría de la probabilidad permite la descripción de la incertidumbre asociada a los sistemas estocásticos naturales y el ajuste de modelos”.

Por un lado, la probabilidad frecuentista se basa en la idea de que la probabilidad de un evento se puede determinar observando la frecuencia con la que dicho evento ocurre en una serie de repeticiones de un experimento. Por ejemplo: si lanzamos un dado honesto (entendamos por “honesto”, no cargado), la probabilidad -en términos frecuentistas- de obtener un “4” sería igual a $1/6$, ya que se espera que aparezca, en promedio, en uno de cada seis lanzamientos. La probabilidad frecuentista se basa, entonces, en la noción de repetición del experimento y la recopilación de datos empíricos. Y la podemos pensar como una probabilidad a “priori” es decir, cuando el evento aún no ha ocurrido.

Por el otro, la probabilidad bayesiana, fundamentada en el teorema de Bayes²⁰, se utiliza para actualizar la probabilidad de un evento en función de la información disponible. Es decir, se basa en las preexistentes probabilidades subjetivas o personales a un evento antes de que se realice el experimento y la actualización de esas probabilidades a medida que se obtienen nuevos datos. Por ejemplo: un editor nos dice que, en un libro de sólo tres capítulos, el 85% de las páginas del primer capítulo no tienen errores, lo mismo ocurre con el 90% de las del segundo y el 95% de las del tercero. Se sabe, además, que el primer capítulo tiene 125 páginas, el segundo tiene 150 y el tercero, 175. Si elegimos una página al azar y comprobamos que no tiene error alguno, ¿cuál es la probabilidad de que, la página escogida, pertenezca al segundo capítulo? Por consiguiente, el teorema de Bayes se utiliza para combinar la información previa (probabilidad a priori) con la nueva evidencia (probabilidad condicional) y, así, obtener la probabilidad buscada. A esta probabilidad la podemos concebir como una probabilidad a “posteriori”, esto es cuando ya ha ocurrido un evento (Barrera-Causil, Sandoval y Sepúlveda-Murillo, 2011).

La principal discrepancia entre ambas miradas radica en su interpretación y en la forma en que se utilizan. La probabilidad frecuentista se basa en la observación empírica y en la sucesiva repetición de un determinado experimento, mientras que la probabilidad bayesiana se basa en las preexistentes probabilidades subjetivas y la actualización de esas probabilidades en función de la evidencia disponible.

En resumen, la probabilidad frecuentista se centra en la frecuencia observada de un evento a partir de repeticiones del experimento, mientras que la probabilidad bayesiana se basa en la actualización de probabilidades subjetivas a medida que se recopila nueva evidencia. Ambos enfoques son

²⁰Thomas Bayes (1702-1761). Matemático, Estadístico y Ministro Presbiteriano británico, creador del Teorema que lleva su nombre.

utilizados en diferentes contextos y tienen sus propias ventajas y limitaciones.

Ventajas y Limitaciones

Antes de presentar algunas de las diferencias (y, por ende, ventajas y limitaciones) entre probabilidades frecuentistas y bayesianas, y de acuerdo con Moral-Benito (2010), es menester aclarar los aspectos comunes o “puntos de encuentro” que tienen los dos modelos. Por un lado, en ambos casos se utilizan modelos con parámetros desconocidos para caracterizar el mundo real y, por el otro, ambos enfoques requieren de la recolección de datos (toma de muestras) como base para la estimación de dichos parámetros desconocidos.

Ahora bien, la probabilidad frecuentista es un enfoque para calcular la probabilidad basada en la frecuencia relativa de un evento en un número grande de repeticiones de un experimento. A continuación (Rolleri, 2004), se presentan algunas ventajas y limitaciones asociadas con este enfoque:

1. Ventajas

- a. Fundamento empírico: se basa en la observación y recopilación de datos empíricos. Esto la hace adecuada para situaciones en las que se realizan repeticiones de un experimento y se recopilan datos reales.
- b. Intuición natural: se alinea con nuestra intuición natural sobre la probabilidad. Si un evento ocurre con mayor frecuencia en repeticiones de un experimento, intuitivamente esperamos que su probabilidad sea mayor.
- c. Validez en el largo plazo: se ajusta a la idea de que la probabilidad de un evento se estabiliza a medida que se realiza un número suficientemente grande de repeticiones del experimento. A medida

que aumenta el número de repeticiones, la frecuencia relativa se acerca a la probabilidad real del evento.

2. Limitaciones

- a. Requiere repeticiones: solo es aplicable cuando es posible repetir un experimento en condiciones similares y recopilar datos sobre la frecuencia relativa del evento. En muchos casos, no es posible realizar repeticiones y, por lo tanto, este enfoque no es aplicable.
- b. No proporciona información sobre eventos únicos: no puede asignar una probabilidad a eventos únicos o eventos que no se pueden repetir. No puede responder preguntas como “¿cuál es la probabilidad de que mañana ocurra, puntualmente, un evento determinado?” ya que no se pueden realizar repeticiones controladas del evento.
- c. Dependencia del tamaño de la muestra: la precisión de la estimación de la probabilidad frecuentista depende del tamaño de la muestra. Si el tamaño de la muestra es pequeño, la estimación puede ser menos confiable y más sensible a la alternativa.

Lo dicho anteriormente respecto al enfoque frecuentista, y según lo expresado por Penna et al. (2016, pp.9-10), lo podemos sintetizar en la siguiente expresión:

$$\frac{f_A}{n} = P(A)$$

Donde A es un evento cualquiera, f_A es la frecuencia absoluta de ocurrencia del evento A , $\frac{f_A}{n}$ la frecuencia relativa de ocurrencia del evento A y $P(A)$ la probabilidad de ocurrencia del evento A .

Por su parte, la probabilidad bayesiana -de acuerdo a lo expresado por Quero-Virla (2020)- es un enfoque para calcular la probabilidad (como ya se dijo) basado en el uso del teorema de Bayes y la actualización de creencias a

medida que se obtiene nueva información (datos). Seguidamente, se presentan algunas ventajas y limitaciones asociadas con esta mirada:

1. Ventajas

- a. Actualización de creencias: permite actualizar las creencias iniciales a medida que se obtienen nuevos datos. Esto es especialmente útil en situaciones en las que se tienen datos previos y se desea incorporar nueva información para obtener una estimación más precisa de la probabilidad.
- b. Flexibilidad: admite incorporar conocimientos previos o información experta en el cálculo de la probabilidad. Esto puede ser especialmente útil cuando hay información previa, veraz y confiable, que puede ayudar a mejorar las estimaciones de probabilidad.
- c. Manejo de incertidumbre: proporciona un marco para expresar y manejar la incertidumbre. Permite representar la probabilidad como una distribución y actualizarla a medida que se obtiene nueva evidencia, lo que facilita la toma de decisiones en situaciones inciertas.

2. Limitaciones

- a. Dependencia de las creencias iniciales: está influenciada por las creencias iniciales. Si las creencias iniciales están sesgadas o son incorrectas, pueden afectar las estimaciones posteriores de probabilidad.
- b. Necesidad de especificar distribuciones previas: para aplicar esta probabilidad, es necesario especificar una distribución previa sobre las probabilidades. La elección de la distribución previa puede ser subjetiva y, por ende, influir en los resultados.
- c. Cálculos intensivos: a medida que se incorpora más información y se actualizan las creencias, los cálculos pueden volverse computacionalmente intensivos, especialmente cuando se tienen

múltiples variables y datos complejos. Esto puede dificultar, en algunos casos, la aplicación práctica de la probabilidad bayesiana.

Según lo expresado por Penna et al. (2016, p.9), lo dicho previamente en relación al enfoque bayesiano -en su versión simplificada y para no entrar en tecnicismos- se sintetiza de la siguiente forma:

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

Siendo A y B eventos cualesquiera, $P(A|B)$ es la probabilidad de ocurrencia del evento A condicionada a la ocurrencia del evento B , $P(B|A)$ la probabilidad de ocurrencia del evento B sabiendo que ocurrió el evento A , $P(A)$ la probabilidad de ocurrencia del evento A y $P(B)$ la probabilidad de ocurrencia del evento B (también conocida, esta última probabilidad, como el Teorema de la Probabilidad Total).

Conclusión

Como cierre del presente trabajo, y a modo de resumen, les dejamos las ventajas y limitaciones de cada uno de los enfoques presentados:

- La probabilidad frecuentista tiene la ventaja de estar fundamentada en datos empíricos y alinearse con nuestra intuición sobre la probabilidad. Sin embargo, su aplicación está limitada a situaciones en las que se puedan realizar repeticiones y recopilar datos sobre la frecuencia relativa del evento. Además, no proporciona información sobre eventos únicos y depende del tamaño de la muestra para una estimación precisa de la probabilidad.
- La probabilidad bayesiana tiene la ventaja de permitir la actualización de creencias y manejar la incertidumbre de manera flexible. Sin embargo, su aplicación está condicionada por la veracidad de las creencias iniciales y la necesidad de especificar distribuciones previas.

Además, los cálculos pueden volverse computacionalmente intensivos en situaciones complejas.

A modo de colofón, y luego de lo expresado anteriormente, dejamos a ustedes, investigadoras e investigadores, la elección respecto a la “mirada” que van a utilizar, tanto en Ciencias Sociales como en Ciencias Humanas, de acuerdo a las características de vuestras investigaciones.

Donde, y para tener muy en cuenta, acorde a lo expresado por Kadane y Lazar (2004), “la selección de modelos es una parte importante de cualquier análisis estadístico y, de hecho, es fundamental para la búsqueda de la ciencia en general”, y es esto lo que nos lleva a pensar cuál es el “mejor” modelo a ser utilizado (p.279)

Dicho de otra manera, son ustedes quienes tendrán que decidir qué enfoque o mirada usar para, a partir de muestras, tomar decisiones válidas y confiables respecto, y como ya se mencionó, a parámetros desconocidos en relación a la población con la cual trabajan.

“¡Datos, datos, datos!”–exclamaba con impaciencia–. “¡No puedo hacer ladrillos sin arcilla!”...profería Sherlock Holmes (Doyle, 1892, p.16). Por ende, las decisiones que ustedes tomen, son las que construirán el conocimiento, ergo, cómo configuremos esos “ladrillos” es lo que nos dará un resultado más robusto o bien, un castillo de naipes. En otras palabras, acorde a los objetivos presentes en su investigación, será el modelo que elijan la encrucijada, por cierto, nada sencilla, que deberán resolver...

Bibliografía

Barrera-Causil, C.J.; Sandoval, J.D.J. y Sepúlveda-Murillo, F.H. (2011). Estimación por intervalos de probabilidad a posteriori para la proporción de estudiantes universitarios desertores. *TecnoLógicas*, (27), 75-87. <http://www.scielo.org.co/pdf/teclo/n27/n27a05.pdf>

- Doyle, A.I.C. (1892). "El misterio de Copper Beeches". [http://bibliotecadigital.ilce.edu.mx/Colecciones/ObrasClasicas / docs/MisterioCopper_Doyle.pdf](http://bibliotecadigital.ilce.edu.mx/Colecciones/ObrasClasicas/docs/MisterioCopper_Doyle.pdf)
- Kadane, J.B. y Lazar, N.A. (2004). "Methods and criteria for model selection". *Journal of the American Statistical Association*, 99: 279-290.
- Moral-Benito, E. (2010). "Econometría: ¿Bayesiano o Frecuentista?". *Sintetia*. <https://www.sintetia.com/econometria-bayesiano-o-frecuentista/>
- Penna, F.O.; Esteva, G.C.; Cobos, O.H.; Ulagnero, C.A. y Vázquez Ferrero, S.M. (2016). "Fórmulas y Tablas II (Para cursos de Estadística Básica)". Nueva Editorial Universitaria, UNSL: San Luis.
- Quero-Virla, M.E. (2020). "O Bayesianismo: Do casuístico ao histórico", *SAPIENTIAE: Revista de Ciências Sociais, Humanas e Engenharias*, vol. 6, núm. 1, pp. 124-130.
- Rolleri, J.L. (2004). "La interpretación frecuentista de la probabilidad: su inaplicabilidad a sucesos singulares". *Signos Filosóficos*, vol. VI, núm. 11, enero-junio, 2004, pp. 159-171. Universidad Autónoma Metropolitana Unidad Iztapalapa: Distrito Federal, México.
- Ruiz-Benito, P.; Andivia, E.; Archambeau, J.; Astigarraga, J.; Barrientos, R.; Cruz-Alonso, V.; Florencio, M.; Gómez, D.; Martínez-Baroja, L.; Quiles, P.; Rohrer, Z.; Santos, A.M.C.; Velado, E.; Villén-Pérez, I. y Morales-Castilla, I. (2018). "Ventajas de la estadística bayesiana frente a la frecuentista: ¿por qué nos resistimos a usarla?". *Ecosistemas*, 27(2): 136-139.
- Russell, B.A.W. (1989). *Introduzione alla filosofia matematica*. Grandi Tascabili Economici Newton: Roma.



Recibido: 27/06/2023

Aceptado: 04/10/2023

Cómo citar este artículo:

Penna, F., Cobos, H., Ferrero Vázquez, S. y Penna, A. (2023). Investigación en Ciencias Sociales y Humanas: ¿Frecuentista o Bayesiana?. *RevID, Revista de Investigación y Disciplinas*, Número 9, San Luis, p 70-80

