

Mauricio Molina Delgado

Bayes y el círculo de la probabilidad

Abstract: *In this article we discuss the debate about probability between frequentalism and bayesianism approaches. Thomas Bayes' original work and other early XVIII century proposals about probability are confronted with the position called bayesianism. We present arguments supporting a subjectivist notion of probability, which considers probability as degree of belief in contrast with the frequentalist approach, which defines the concept in terms of the limit of relative frequencies when the number of cases goes to infinity.*

Key words: *Bayes. Probability. Hazard. Determinism. Hermeneutics.*

Resumen: *En este artículo se aborda la polémica entre frecuentistas y bayesianos respecto a la naturaleza de la probabilidad. Al respecto se considera la figura de Thomas Bayes y las nociones tempranas de probabilidad en el siglo XVIII, contrastando su propuesta con el enfoque que posteriormente se denomina bayesiano. Se dan argumentos a favor de la noción subjetivista de probabilidad, la cual considera que ésta representa un grado de creencia, en contraste con la versión frecuentista que define este concepto en términos de una frecuencia relativa cuando el número de casos tiende al infinito.*

Palabras clave: *Bayes. Probabilidad. Azar. Determinismo. Hermenéutica.*

Sobre la noción clásica de probabilidad

Normalmente se entiende por probabilidad de un evento A al límite llevado al infinito de la razón entre el número de casos donde se obtiene el evento A sobre el número total de ensayos. Es esta la noción que podríamos denominar enfoque frecuentista, el cual supone que las probabilidades son valores fijos que se ven reflejados empíricamente al repetir muchas veces un experimento en condiciones similares. Dicha interpretación frecuentista de la probabilidad no es la única posible, de hecho inicialmente se adoptó un concepto de probabilidad como una medida del grado de creencia en una situación de ignorancia parcial y esta noción expresada ya en Laplace fue aceptada por la mayoría de los autores hasta el siglo XIX (Nagel, 1965). Esta última idea ha sido tradicionalmente adjudicada a Thomas Bayes cuyo trabajo, publicado de forma póstuma, ha sido interpretado en el sentido de que los parámetros de una distribución de probabilidad no son fijos sino variables (Bayes, 1763/2003).

En general podríamos señalar 3 posiciones distintas sobre el concepto de probabilidad¹: a) nociones que recurren a resultados empíricos y asumen que las frecuencias relativas son un reflejo del valor de probabilidad, b) nociones que consideran a la probabilidad como un grado de creencia en situaciones de ignorancia parcial y que recurren al supuesto de eventos equiprobables, y c) nociones que consideran a la probabilidad como un grado de creencia en situaciones

de ignorancia parcial pero no asumen la equiprobabilidad de los eventos sino que recurren al conocimiento subjetivo para obtener valores de probabilidad. Se suele denominar definición clásica de probabilidad al segundo grupo de enfoques (b), definición frecuentista o estadística al primer grupo (a) y definición personalista o bayesiana al tercero (c).

Para los efectos que aquí interesan, existen dos formas básicas en las que podemos entender la probabilidad: una forma objetivista basada en la idea de límite al infinito de la frecuencia y una forma personalista que considera la probabilidad como un grado de creencia. El propósito de este artículo es presentar argumentos que muestren la superioridad de la definición personalista o bayesiana respecto al enfoque objetivista.

Sobre la naturaleza de la estadística

La discusión precedente se restringe únicamente al tema de la probabilidad en cuanto a su definición y cálculo, sin embargo el campo de la estadística trasciende al de la probabilidad si bien está construido a partir de ella.

A partir de la noción frecuentista de la probabilidad se ha desarrollado el llamado enfoque clásico de la estadística², el cual corresponde en realidad a una difícil alianza entre dos propuestas: la de Fisher y la de la llamada escuela de teoría de decisión Neyman-Pearson-Wald (Efron, 1986)³. Ambos enfoques surgen en los años 30 del siglo XX y una discusión profunda de lo mismos está más allá de las posibilidades de este artículo. Digamos simplemente que el enfoque clásico de la estadística supone que la probabilidad de ocurrencia de cierto evento X se puede modelar con una determinada función $f(X)$ la cual depende de un parámetro o valor fijo que es en principio desconocido. Así, por ejemplo, el ensayo de lanzar una moneda al aire cuenta con una probabilidad de que ésta caiga en escudo, la cual puede ser expresada por una función que denominamos Bernoulli con parámetro p . Si asumimos que ambas caras de la moneda tienen idéntica probabilidad de ocurrir, el parámetro p debería ser igual a 0.5. Es también posible que en

realidad dicho parámetro asuma cualquier valor entre 0 y 1, pero según el enfoque clásico dicho valor es necesariamente fijo aún cuando pueda ser desconocido.

El enfoque bayesiano de probabilidad

La obra fundamental de Bayes (1763/2003) es un artículo póstumo en el que trata de la determinación de la probabilidad de que el parámetro p de una distribución Bernoulli se encuentre entre dos valores. Esta formulación contrasta con el enfoque tradicional en cuanto considera los parámetros no como cantidades fijas sino como variables a cuyos posibles valores se asocian a su vez valores de probabilidad.

Algunos autores consideran que antes del surgimiento del enfoque clásico de la estadística esta disciplina era fundamentalmente bayesiana (Efron, 1986). Esta interpretación es sin embargo difícil de aceptar. Si nos remontamos al siglo XVII más bien podríamos hablar de que algunos desarrollos de la estadística tenían un tinte bayesiano, como sucede con el trabajo de Laplace y Gauss, pero la verdad es que aun en nuestros días se utilizan herramientas bayesianas desde una perspectiva clásica. De hecho, dada la oscuridad de su artículo es dudoso que podamos decir que Bayes (1763/2003) mismo fuera bayesiano. Lo cierto es que contemporáneamente surgen una serie de posiciones basadas en un enfoque subjetivista de las probabilidades⁴.

Se ha interpretado que Bayes (1763/2003) considera que en ausencia de ignorancia sobre la distribución del parámetro se debe considerar que todos los valores posibles del mismo tienen igual probabilidad de ocurrir. Lo importante es que Bayes, tal como se había adelantado, considera al parámetro no como un valor fijo (como aparece en la formulación clásica) sino como variable. En segundo lugar, el teorema de Bayes propone que la determinación de la probabilidad depende tanto del conocimiento previo como de los datos mismos.

Ya en el siglo XX la discusión entre el enfoque bayesiano y el enfoque clásico parece tener algunos de los rasgos descritos por Kuhn

(19870) respecto de las revoluciones copernicanas (Smith, 1986). Los defensores del bayesianismo argumentan que la supremacía de la definición de probabilidad frecuentista adolece de fallas importantes (Gill, 2002; Press, 2003). Se han señalado algunas paradojas que surgen de ese enfoque debidas a la ruptura del llamado principio de verosimilitud o la tendencia al uso irreflexivo de significancias del 1% y el 5% (al respecto ver Cohen, 1994). Según esta interpretación, la razón de que el paradigma clásico se mantuviera durante casi 100 años se debe a las dificultades matemáticas del enfoque bayesiano, pero al desarrollarse recientemente métodos de simulación de Monte Carlo y herramientas computacionales para su implementación, estos finalmente vendrían a jugar el papel que tuvo el telescopio en la instauración del heliocentrismo.

En el presente artículo se presentan dos nuevos argumentos, uno centrado en la concepción de probabilidad (en adelante argumento del mundo determinístico) y otro en la utilidad práctica de la misma respecto de los procesos de interpretación de la realidad (argumento hermenéutico).

Argumento del mundo determinístico

Considérese una sucesión de n experimentos de Bernoulli (es decir, aquel donde se produce al azar uno de 2 posibles eventos que arbitrariamente denominamos éxito y fracaso) como sería el lanzamiento de una moneda al aire. Para $n = 6$ podríamos tener los resultados presentados en la siguiente tabla, el resultado del experimento se representa con 1 cuando sale escudo (arbitrariamente lo consideraremos el éxito) y 0 cuando sale corona (el fracaso).

| | | | | | | | |
|-----------------------|---|---|---|---|---|---|---|
| Número de lanzamiento | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Resultado | 1 | 0 | 0 | 0 | 1 | 1 | ? |

Supongamos por un momento que en realidad los resultados del lanzamiento de la moneda no se dan al azar sino que existe una ley que los gobierna. De este modo, el proceso sería determinístico, existiendo un patrón que permitiría

predecir el resultado de la siguiente tirada del dado. Por ejemplo, supongamos que el siguiente conjunto de reglas se cumple:

- R1) Si los últimos 3 experimentos (lanzamientos de la moneda) al ser sumados producen un número impar el resultado siguiente será idéntico al del último experimento.
- R2) Si los últimos 3 experimentos (lanzamientos de la moneda) al ser sumados producen un número par o 0, el resultado siguiente será diferente al del último experimento (es decir, si el último experimento dio 1 el siguiente será 0 y viceversa).

Así, los 3 primeros resultados suman 1 (impar) y el tercero es 1 de modo que el cuarto deberá ser 1 (por R1); los resultados de los experimentos 2, 3 y 4 suman 0 y el cuarto es 0 de modo que el quinto deberá ser 1 (por R2); y finalmente los resultados de los experimentos 3, 4 y 5 suman 1 por lo que el sexto deberá ser 1 (impar) ya que el resultado del quinto experimento es 1 (por R1). Podemos obviar el hecho de que los primeros 3 resultados no pueden ser determinados por las reglas R1 y R2 suponiendo simplemente que los seis ensayos pertenecen a una serie más grande de la que no tenemos conocimiento pero que igualmente determinarían los resultados de los primeros resultados. Agreguemos que aunque el proceso es determinístico, según hemos supuesto en nuestro experimento mental, el sujeto que lanza las monedas, al que llamaremos en adelante Gombaudo, desconoce el patrón que siguen los datos. Ahora bien, si Gombaudo decide lanzar por séptima vez la moneda, nosotros podemos predecir que el resultado debe ser 0 puesto que la suma de los últimos 3 resultados da 2 y el último resultado fue 1 (R2). Desde nuestra perspectiva, como conocedores del proceso determinístico detrás del experimento, no tiene sentido considerar una probabilidad, pero para Gombaudo el asunto es distinto. Si él mantiene una noción personalista de probabilidad le aceptaremos que nos hable en términos probabilísticas, porque a pesar de que el proceso es determinístico su noción de probabilidad no encuentra tropiezo en ello, ya que al decirnos por ejemplo que él espera un escudo con una probabilidad de 0.5 sencillamente está

expresándonos su grado de creencia en un contexto de ignorancia parcial. Por el contrario, si Gombaud mantiene un compromiso frecuentista al darnos un enunciado probabilista estaría cometiendo un error. En este caso la probabilidad de que se obtenga un escudo en la tirada número 7 sería calculable únicamente si fuera posible repetir de forma idéntica las condiciones. Gombaud no lo sabría, pero para nosotros como genios omniscientes las condiciones relevantes serían las propias de las reglas R1 y R2, es decir la definición frecuentista de probabilidad nos diría que la probabilidad de escudo (1) en la siguiente tirada debe ser igual a la proporción de veces que la suma de los tres experimentos anteriores es igual a 2 y el último es igual a 1, en cuyo caso las reglas R1 y R2 nos garantizan que esa proporción es necesariamente 0. Gombaud no tiene acceso a este resultado, por lo que casi cualquier número que él nos dé tendrá poco que ver con la situación objetiva de la probabilidad.

Modifiquemos ahora un poco nuestro experimento. Supongamos que finalmente Gombaud lanza la moneda y obtiene escudo (1). El resultado echa por tierra nuestra suposición de que R1 y R2 gobiernan el proceso, sin embargo esto no garantiza que efectivamente la secuencia de resultados sea aleatoria. De hecho, una nueva serie de reglas podrían dar cuenta del patrón, considérese por ejemplo las reglas R3 a R5, donde el símbolo X representa alguno de los valores posibles (0 o 1) y X' su complemento (es decir que si $X = 0$ entonces $X' = 1$ y si $X=1$ luego $X'=0$)

- R3) Cuando aparezca una sucesión de resultados X'X el siguiente resultado será X.
- R4) Cuando aparezca una sucesión de resultados X'XX el siguiente resultado será X.
- R5) Cuando aparezca una sucesión de resultados X'XXX el siguiente resultado será X'.

En otras palabras las R3 a R5 reglas generan rachas de 1, 2 ó 3 resultados idénticos seguidos, pero no rachas de 4 ó más. Nuevamente podemos asumir que el resultado del primer ensayo depende de otros resultados no conocidos, y a partir de allí se puede demostrar que los siguientes son correctamente predichos por el nuevo sistema de reglas, incluyendo el séptimo ensayo. Como

puede verse, aunque las reglas R1 y R2 son refutadas a la séptima tirada de la moneda, otro grupo de reglas sí puede explicar los siete resultados. El punto es que siempre es posible encontrar un conjunto de reglas que explique cualquier grupo de resultados⁵, y aunque Gombaud puede desconocer dichas reglas siempre es posible que ellas existan y teóricamente siempre podríamos pensar que alguien posee el conocimiento de las mismas. Así, Gombaud desconocía en nuestro ejemplo las reglas correctas, nosotros planteamos un conjunto y ese resultó ser incorrecto, pero siempre una tercera persona podría poseer el conocimiento correcto. Esto puede utilizarse como argumento para cuestionar la posibilidad del azar, pero desde la perspectiva que aquí nos interesa el punto es que, independientemente de la existencia o no de un conjunto de reglas determinísticas (universalmente o localmente respecto al problema específico de las monedas), Gombaud es incapaz de predecir la siguiente tirada de la moneda y que nosotros también fracasamos en la predicción. En este sentido las predicciones (tanto la de Gombaud como la nuestra) son de naturaleza probabilística siempre que adoptemos una definición personalista de probabilidad. Sin embargo si adoptamos una definición frecuentista resulta que Gombaud no es capaz de emitir ningún enunciado probabilística ya que no es capaz de reproducir las condiciones que las reglas proponen. En un mundo determinístico la posición frecuentista pierde sentido, pero la definición personalista no tiene ningún compromiso en este sentido, ya que lo único que requiere es la ignorancia relativa del sujeto. Puesto que no tenemos forma de afirmar o refutar que el proceso (o el universo mismo) sea de naturaleza determinística, el planteamiento personalista es claramente superior.

Argumento hermenéutico

Gadamer (2003) ha recalcado la idea de que los procesos de comprensión se llevan a cabo siempre a partir de preconcepciones (nunca se parte de conocimiento cero) las cuales eventualmente pueden ser revisadas y modificadas. El

teorema de Bayes (1763/2003) puede ser descrito en el sentido de que la probabilidad a posteriori de un parámetro es función tanto de la probabilidad a priori (es decir la creencia del sujeto, la cual vendría siendo un indicador del prejuicio del sujeto o de su conocimiento en un momento dado) como de la verosimilitud de los datos (es decir la evidencia recolectada). Esto implica que las probabilidades que previamente asume el sujeto se actualizan a partir de la revisión empírica. El teorema de Bayes también es coherente con la idea de un proceso ilimitado de actualizaciones del conocimiento (Gill, 2002). Por otra parte, el llamado círculo hermenéutico de Gadamer parece contar con algún apoyo en los desarrollos en psicología que muestran el papel de esquemas (Brewer y Treyns, 1981) y estereotipos (Lambert y Chasteen, 2004)⁶.

Independientemente de que aceptemos o no la hermenéutica gadameriana, lo relevante es que ésta representa al menos una explicación plausible de los procesos de comprensión. Al ser coherente con el enfoque bayesiano le da a éste algún viso de validez psicológica. Por el contrario, definitivamente el enfoque frecuentista carece en muchos casos de dicha validez. Especialmente cuando la idea de replicación de los experimentos es imposible o poco creíble.

Conclusión

Junto a los argumentos hasta ahora esgrimidos a favor de la definición personalista de probabilidad, los cuales tradicionalmente han sido de naturaleza más bien estadística, se ha mostrado que la definición personalista de probabilidad tiene ventajas conceptuales al no requerir del supuesto de que el universo (o al menos la faceta que en un momento dado es analizada desde un punto de vista probabilístico) no corresponde a procesos determinísticos.

Los defensores del frecuentismo han reconocido la superioridad en el peso filosófico del bayesianismo subjetivo (Efron, 1986) sin embargo han argumentado a favor de las ventajas prácticas de su postura, específicamente la sencillez de los procedimientos y la necesidad de una

posición objetivista. La discusión precedente sin embargo, ha mostrado que el enfoque personalista cuenta con una ventaja práctica importante ya que representa o al menos se asemeja a mecanismos psicológicos plausibles. En este sentido la exigencia objetivista, en el sentido planteado por los frecuentistas, tampoco parece ser coherente con la dinámica histórica del desarrollo científico (Kuhn, 1970). De hecho, las bondades prácticas del planteamiento subjetivista se han visto reflejadas en el hecho de que efectivamente el teorema de Bayes ha sido aplicado con éxito en algunos campos de la psicología, tales como los referentes a la toma de decisiones.

Por último, habría que señalar que esto no implica que el enfoque frecuentista sea totalmente inútil. De hecho, de manera análoga a la situación que se presenta entre la física newtoniana y la física relativista, el frecuentismo es en realidad un caso particular de la formulación bayesiana: específicamente el caso donde el conocimiento previo no es particularmente informativo.

Notas

1. No necesariamente un concepto de probabilidad debería derivar en un cálculo de probabilidad, sin embargo aquí se incluyen únicamente los casos donde se propone explícitamente un cálculo. Ángel (1965) discute una mayor gama de posibles interpretaciones de la probabilidad.
2. Debe aquí diferenciarse dicho enfoque clásico de la definición clásica de probabilidad anteriormente presentada. Como se dijo, el enfoque clásico tiene un compromiso con la definición frecuentista de probabilidad, no con la definición llamada clásica.
3. El aporte de Fisher se refiere a lo que se ha denominado estimación máximo verosímil, lo cual corresponde a un modo de estimar el valor de los parámetros fijos; mientras que la escuela de Neyman-Pearson-Wald han desarrollado aspectos como el de la prueba de hipótesis.
4. Debe tenerse en cuenta que existe también una línea bayesiana objetivista, sin embargo en la presente discusión interesa únicamente la versión subjetivista del bayesianismo.

5. Goodman (1984) apunta precisamente a este hecho. A posteriori siempre podrá encontrarse una serie de reglas que den cuenta de los datos, pero también es posible encontrar un polinomio que lo haga (Martínez, 2008).
6. En realidad estos desarrollos muestran más bien la persistencia de estas estructuras previas incluso cuando se enfrentan con la evidencia. Sin embargo la ruptura de esos esquemas es según estos estudios una posibilidad, aunque requiere ciertamente de un esfuerzo cognitivo extra por parte del sujeto.

Bibliografía

- Bayes, T. (1763-2003) “ An essay towards solving a problem in the doctrine of chances”. In S. J. Press & C. Siddartha (Eds.), *Bayesian statistics: Principles, models and applications*. (pp. 423-448). Hoboken, NJ: Wiley.
- Cohen, J. (1994) The earth is round ($p < .05$). *American Psychologist*, *49*(12), 997-1003.
- Gadamer, H.G. (2003) *Verdad y método*. Salamanca: Sígueme.
- Gill, J. (2002) *Bayesian methods: A social and behavioral sciences approach*. New York: Chapman and Hall.
- Goodman, N. (1976-1984) *Languages of art. An approach to a theory of symbols*. Indiana: Hackett.
- Kuhn, T. (1970-1996) *The scientific revolutions*. Chicago: Chicago University Press.
- Martínez, G. (2008, noviembre). *Psicología, series lógicas y cálculo probabilística. Encuentro de docentes e investigadores de Estadística en Psicología*. Buenos Aires, Argentina.
- Nagel, E. (1965) *Principles of the theory of probability*. Chicago: Chicago University Press.
- Press, S.J. (2003) *Subjective and objective bayesian statistics: Principles, models and applications*. Hoboken, NJ: Wiley.