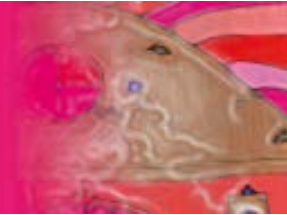




## INVESTIGACIÓN EN PSIQUIATRÍA



### Utilidad del método de Análisis Estadístico Bayesiano en psiquiatría

Autor: Daniel Serrani (psiquiatra, geriatra universitario, miembro del capítulo de Investigación en Psiquiatría de APSA, Dr. en Psicología, Mg en Salud Mental, Dpl. Políticas Públicas en Salud Mental) correo electrónico: [danielserraniazcurra@gmail.com](mailto:danielserraniazcurra@gmail.com), [daniel.serrani@unr.edu.ar](mailto:daniel.serrani@unr.edu.ar)

El análisis bayesiano es un método de [inferencia estadística](#) de una [probabilidad condicional](#) que difiere de los métodos convencionales basados en frecuencias como la prueba de hipótesis nula. Una probabilidad condicional (que el evento A ocurra a condición que ocurra B), se basa en suposiciones a priori que el método bayesiano permite actualizar a medida que se acumula nueva evidencia. Si ambos sucesos son posibles (probabilidad no nula) la probabilidad condicional de A dado B se define como:

$$Pr(A/B) = \frac{Pr(A \cap B)}{Pr(B)} \quad (1).$$

Si  $A_1, A_2, \dots, A_k$  fueran  $k$  sucesos mutuamente excluyentes, uno de los cuales debe ocurrir necesariamente; la ley de la probabilidad total establece que:  $Pr(B) = \sum_{i=1}^k Pr(B/A_i)P(A_i)$  (2)

Tomando el suceso  $A_j$  en lugar de  $A$  en la fórmula y aplicando al denominador la mencionada ley, tenemos:  $Pr(A_j/B) = \frac{Pr(B/A_j)P(A_j)}{\sum_{i=1}^k Pr(B/A_i)P(A_i)}$  (3).

La hipótesis nula determina el valor único de un parámetro poblacional ajustado a los datos y también establece un rango de valores posibles dentro del intervalo de confianza, lo cual indica la precisión de la medición. La inferencia bayesiana estima parámetros potenciales asignándoles diferentes pesos en una distribución de probabilidades que el investigador evalúa. Ante preguntas tales como “¿qué certeza hay que el tratamiento A sea más eficaz que el B?” o “¿qué nivel de evidencia hay de una asociación no trivial entre dos variables?” el método bayesiano ofrece

respuestas en forma de distribución de probabilidades (por ejemplo: 75% o 92%) basándose en los conocimientos previos y los datos observados permitiendo derivar interpretaciones acerca de los modelos de parámetros de interés, como por ejemplo: 75% de probabilidad de que  $\beta > 0$ . La utilidad es que permite analizar la robustez de evidencia tanto para la ausencia como la presencia de efectos. Una distribución de probabilidades alrededor de 3 o de 89 representa evidencia de un tamaño de efecto pequeño o grande, respectivamente.

Otras ventajas son las siguientes:

1. Funciona bien con muestras pequeñas, no depende de las propiedades asintóticas de grandes muestras, y ofrece evidencia graduada. Por ejemplo, supongamos que en un estudio pequeño encontramos una probabilidad para una diferencia entre grupos del 72%, que aún estando debajo del umbral de una diferencia clínicamente significativa, permite derivar implicaciones clínicas.
2. Permite incorporar datos en forma secuencial actualizando los análisis hasta obtener un nivel adecuado de precisión en los parámetros sin tener que limitarnos por la necesidad de un cálculo previo del tamaño muestral necesario. Esto es especialmente útil cuando se cuenta con muestras pequeñas, recursos limitados o poco tiempo disponible.
3. Se puede escoger una distribución de probabilidades para los parámetros de interés previa al estudio (priors) y luego del mismo evaluar si los resultados se ajustan al modelo previsto. Las probabilidades previas contienen de modo transparente nuestros supuestos sobre los valores de los parámetros. Por ejemplo, en estudios de regresión logística con predictores y resultados estandarizados, sabemos casi con certeza que rara vez encontraremos coeficientes  $> \pm 5.0$ . Esta información podría ser incluida en el análisis, y la distribución posterior debería aclarar que hemos aprendido agrupando y analizando los datos.

Fotografía Dra. Miriam Monczor, Mar del Plata, 2024

¿Cómo iniciarse en el uso de métodos bayesianos? Si bien puede parecer complejo al inicio, hay varios recursos (1-5) y programas estadísticos: paquetes en python (pymc3, PyStan, Arviz, bayesian toolkit en <https://github.com/bat/bat>), y paquetes en R (rstan, BRMS, BayesFactor).

Pasos en la aplicación de un programa bayesiano:

1. Definir los parámetros del modelo y elegir la probabilidad previa (“prior”) de los datos basados en las suposiciones y conocimientos previos. Podemos hacer regresión, suponiendo variables continuas distribuidos normalmente o emplear distribución t de Student con colas robustas o podemos modelar regresiones sobre variables ordinales, categóricas o dicotómicas.
2. Considerar los “prior” necesarios para el modelo (coeficiente de regresión, error de varianza, etc.) o bien realizando un chequeo predictivo previo con simulaciones para esbozar distribuciones de probabilidades condicionales previas, y a partir de ellas estimar la verosimilitud de los valores de resultados, a través de múltiples iteraciones.
3. Ajustar los datos con el producto de nuestras probabilidades previas y la verosimilitud de datos de acuerdo al [teorema bayesiano](#), para derivar la distribución posterior, sea analíticamente o aproximándose con métodos de cadena de Markov Montecarlo (MCMC).
4. Verificar que la aplicación del MCMC fue satisfactoria a través del análisis de las cadenas del algoritmo que deben converger al mismo resultado.

5. Comprobar el ajuste del modelo a los datos con un chequeo predictivo posterior de la distribución de la variable dependiente implicada en nuestro modelo, que debe reflejar la distribución actual de la variable dependiente, de lo contrario se necesita revisar el modelo.
6. Inspeccionar la distribución posterior estimando la media, y la proporción superior o inferior a algún valor con intervalos creíbles dentro de los cuales debe estar el valor del parámetro verdadero con su correspondiente probabilidad. El rango de los intervalos, depende del nivel de certeza requerido, en caso de un nivel del 95% se requieren muchas muestras de las colas de la distribución posterior con un tamaño efectivo de muestra suficiente, alrededor de 400 para cada parámetro.

La comparación con otros modelos ajustados a los mismos datos permite cuantificar el nivel de evidencia para cada uno. Por ejemplo, contrastar el teorema bayesiano con una regresión para los efectos de un tratamiento con interacción por tiempo, o con una validación cruzada si suponemos que un modelo complejo tiene un ajuste muy robusto, representado particularidades de la muestra antes que rasgos sustanciales del modelo.

Ejemplo de aplicación del teorema de Bayes. Supongamos que queremos calcular la probabilidad que un sujeto que da positivo en un test para determinar riesgo de esquizofrenia, verdaderamente tenga ese riesgo, conociendo la prevalencia del mismo en la población a la que pertenece y los valores de sensibilidad y especificidad de la prueba. Si llamamos  $P$  a la probabilidad a priori que el sujeto tenga el riesgo, y  $Q=1-P$  a su complementaria,  $S$  a la sensibilidad y  $E$  a la especificidad de la prueba  $T$ ; aplicando el teorema de Bayes podemos calcular la probabilidad que el sujeto verdaderamente tenga el riesgo cuando da positivo (valor predictivo positivo de la prueba) y que no tenga el riesgo cuando da negativo (valor predictivo negativo). Reescribiendo la fórmula (3) con los datos actuales tenemos la siguiente:

$$P(Enf/T+) = \frac{P/S}{P*S+Q(1-E)} \text{ y } P(NEnf/T-) = \frac{Q*E}{Q*E+P*(1-S)}$$

Si sabemos que la prevalencia en población del riesgo de esquizofrenia es de 0.2% y que el test que efectuamos tiene una sensibilidad del 98 % y especificidad del 98 % ¿cuál es la probabilidad de que un sujeto que ha resultado positivo tenga verdaderamente riesgo de esquizofrenia? Sustituyendo los valores en la fórmula (3) obtenemos una probabilidad de 0.047, es decir cerca del 5% de los positivos en el test son falsos positivos, ¿cómo una prueba con sensibilidad y especificidad altas parezca en la práctica mala?. El problema

radica en que la prevalencia es baja y si se refiere a la población general quizás no sea aplicable a un sujeto que acude a consulta porque hay otros motivos de sospecha –pertenece a un grupo de riesgo, presenta síntomas específicos...– y entonces no es aplicable la prevalencia de la población general, sino la del subgrupo al que pertenece y en el que la prevalencia (probabilidad a priori) de padecer el riesgo será radicalmente mayor.

Sin embargo los cálculos son válidos si pensamos en la población general, porque queremos plantear un programa de "screening" y hay que considerar el coste social, personal y económico que supone tener un gran número de falsos positivos, frente al beneficio de detectar verdaderos sujetos en riesgo. El modelo de Bayes nos permite como en el ejemplo anterior, modificar la probabilidad conocida de que ocurra un suceso cuando tenemos nueva información al respecto.

#### Bibliografía:

1. [Kruschke JK. Introduction. En: "Doing Bayesian data analysis". A tutorial with R, JAGS, and Stan. EE. UU.: Elsevier; 2015. DOI: 10.1016/C2012-0-00477-2](#)
2. [Indrayan A, Kumar-Malhotra R. "Medical biostatistics". EE. UU.: Chapman & Hall/Chemical Rubber Company Press; 2008.](#)
3. [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](#)
4. [Wagenmakers EJ, Marsman M, Jamil T, Ly A, Verhagen J, Love J, et al. "Bayesian inference for psychology". "Part I: theoretical advantages and practical ramifications" Psychon Bull Rev. 2018;25\(1\):35-57. DOI: 10.3758/s13423-017-1343-3](#)
5. [Etz A, Gronau QF, Dablander F, Edelsbrunner PA, Baribault B. "How to become a Bayesian in eight easy steps: an annotated reading list" Psychon Bull Rev. 2018;25\(1\):219-234. DOI: 10.3758/s13423-017-1317-5](#)