

BIOESTADÍSTICA

Métodos Alternos de Análisis Estadístico en Epidemiología: Frecuentista, Verosimilitud y Bayesiano

ERICK SUÁREZ, PhD; CYNTHIA M. PÉREZ, PhD; JOSUÉ GUZMÁN, PhD.

Una de las críticas a las pruebas de significancia, basadas en los intervalos de confianza y en el valor de p (p -value), es su interpretación utilizando el concepto de repeticiones. Este tipo de interpretación se conoce como la interpretación frecuentista (*frequentist interpretation*), la cual indica una repetición de las muestras de estudio bajo condiciones similares, condición difícil de aplicar en estudios epidemiológicos observacionales. Por lo tanto, se han desarrollado otras alternativas que puedan contestar esta interrogante. Una de estas alternativas es el cálculo de los intervalos de verosimilitud (*likelihood intervals*); otra se obtiene

mediante el cálculo de intervalos bayesianos (*Bayesian intervals*). En este documento presentamos una breve descripción de estos métodos y sus aplicaciones en estudios epidemiológicos. Se espera que con el desarrollo de nuevos programas de computación, las aplicaciones de estos métodos irán en aumento, por lo que los investigadores de la salud tendrán otras alternativas de análisis y evaluación de datos estadísticos.

Palabras clave: Intervalos de confianza, Intervalos de verosimilitud, Intervalos bayesianos, Epidemiología.

Una de las críticas a las pruebas de significancia, basadas en los intervalos de confianza y en el valor de p (p -value), es su interpretación utilizando el concepto de repeticiones. Por ejemplo, un intervalo de confianza al 95% implica que con infinitas repeticiones de las muestras, bajo condiciones similares, un 95% de los límites contienen el verdadero parámetro. Este tipo de interpretación se conoce como la interpretación frecuentista (*frequentist interpretation*), la que es difícil de aplicar en estudios epidemiológicos observacionales. La teoría tradicional de un intervalo de confianza no indica si los límites que se obtienen de un estudio en particular contienen el verdadero parámetro (1). Por lo tanto, se han desarrollado otras alternativas que puedan contestar esta interrogante. Una de estas alternativas es el cálculo de los intervalos de verosimilitud (*likelihood intervals*); otra alternativa es el cálculo de intervalos bayesianos (*Bayesian intervals*).

Estimación por Máxima Verosimilitud

Los métodos para estimar un parámetro por intervalos de verosimilitud dependen de la función de verosimilitud, L . Esta función está determinada por la distribución asociada a la variable aleatoria de estudio y los datos observados. La verosimilitud es una medida de apoyo que se obtiene de los datos observados para un parámetro particular de la distribución de probabilidad. Para una muestra aleatoria, (x_1, x_2, \dots, x_n) la función de verosimilitud de una distribución $f(x|\theta)$, está dada por $L = \prod_{i=1}^n f(x_i|\theta)$.

Existen diferentes medidas de verosimilitud o credibilidad, para diferentes valores del parámetro en términos de probabilidad. Por ejemplo, la función de verosimilitud para una variable aleatoria con distribución binomial, dadas n observaciones independientes, es la siguiente (2):

$$L = \pi^x (1 - \pi)^{n-x}$$

donde: π = probabilidad de ocurrencia del evento de interés en un sujeto; $0 < \pi < 1$.

x = sujetos observados en la muestra con el evento de interés; $x = 0, 1, \dots, n$.

n = sujetos observados en la muestra.

Departamento de Bioestadística y Epidemiología, Escuela Graduada de Salud Pública, Recinto de Ciencias Médicas, Universidad de Puerto Rico.

Dirigir correspondencia a: Dr. Erick Suárez, Departamento de Bioestadísticas y Epidemiología, Escuela Graduada de Salud Pública, Recinto de Ciencias Médicas, Universidad de Puerto Rico, PO Box 365067, San Juan, Puerto Rico, 00936-5067. Teléfono: (787) 758-2525 Ext. 1430, Fax: (787) 764-5831. E-mail: e-suarez@rcmaca.upr.clu.edu

Existe un valor en L para los diferentes valores de π , según los datos observados. El valor particular de π con el valor máximo de L se define como el estimador de máxima verosimilitud (*maximum likelihood estimate, MLE*). En el caso binomial, el MLE es la proporción observada de sujetos con el evento de interés, x/n . Cualquier otro valor diferente de esta proporción tendrá un valor de L menor que el MLE:

$$L_{\pi \neq MLE} \leq L_{MLE}$$

En el caso de una distribución *Poisson*, la función de verosimilitud correspondiente es la siguiente (2):

$$L = \lambda^x \exp(-\lambda Y)$$

donde: x = eventos observados en la muestra; $x = 0, 1, 2, 3, \dots$
 λ = tasa o promedio de eventos de interés por unidad de tiempo
 Y = tiempo total de observación en la muestra
 $\exp(a)$ = representa e^a donde e es la base del logaritmo natural.

En este caso, la estimación de máxima verosimilitud para el parámetro λ se alcanza cuando el parámetro es igual a x/Y .

En un estudio caso-control no pareado, la función de verosimilitud para estimar el *odds ratio, OR*, es la siguiente, utilizando la notación de la Tabla 1:

$$L = \frac{Odds_1^b / (Odds_0 + 1)^{M_0}}{(OR \times Odds_0 + 1)^{M_1} / (OR \times Odds_0)^a}$$

El estimador de máxima verosimilitud para el *OR* es ad/bc .

Para facilitar los cálculos, se acostumbra utilizar el logaritmo natural de L , lo cual resulta en la siguiente expresión (2):

$$b \cdot \ln(Odds_0) - M_0 \cdot \ln(Odds_0 + 1) - M_1 \cdot \ln(OR \cdot Odds_0 + 1) + a \cdot \ln(OR \cdot Odds_0)$$

Los datos de un estudio caso-control no-pareado dirigidos a evaluar la asociación entre una exposición y una enfermedad se pueden resumir en una tabla de contingencia 2x2 de la forma siguiente (3,4):

Tabla 1. Notación para Resumir los Datos de un Estudio Caso-Control No Pareado

Exposición	Casos	Controles	Total
Expuesto	a	b	N_1
No-expuesto	c	d	N_0
Total	M_1	M_0	N

$$Odds_0 = b/d, Odds_1 = a/c$$

$$OR = Odds_1 / Odds_0$$

En el caso de un estudio de cohorte prospectivo o retrospectivo, la función de verosimilitud para estimar el *riesgo relativo, RR*, es la siguiente:

$$L = \frac{[RR \cdot (Y_1/Y_0)]^{D_1}}{[1 + RR \cdot (Y_1/Y_0)]^D}$$

donde:

- λ_i = indica la densidad de la incidencia en el grupo "i" (i = 1 expuestos, i = 0 no-expuestos)
- Y_i = tiempo-persona en el grupo "i"
- D_i = número de sujetos que desarrollaron la condición de estudio en el grupo expuesto
- D = total de sujetos que desarrollaron la condición de estudio

En términos logarítmicos, esta función de verosimilitud puede ser expresada de la forma siguiente:

$$\ln(L) = D_1 \cdot \ln [RR \cdot (Y_1/Y_0)] - D \cdot \ln [1 + RR \cdot (Y_1/Y_0)]$$

El estimador de máxima verosimilitud del riesgo relativo será la razón de incidencias observadas:

$$\frac{D_1 / Y_1}{D_0 / Y_0}$$

En este caso particular, la función L correspondiente está condicionada a que $\lambda_1 \cdot Y_1$ y $\lambda_0 \cdot Y_0$, los casos esperados en ambos grupos, sean constantes y no se afecten por el estudio. Este argumento se cumple principalmente cuando se están realizando estudios con eventos raros, es decir, con incidencias muy bajas (2).

Para determinar la relación proporcional entre la función L de cualquier valor del parámetro, diferente al MLE, y a L del MLE, se puede obtener la razón de verosimilitud (*likelihood ratio, LR*) de la forma siguiente:

$$LR = L_{\pi} / L_{MLE}$$

Para cualquier otro valor del parámetro de estudio (π , λ , OR , RR, \dots), la razón de verosimilitud es menor de uno. Si el valor del parámetro es el MLE, entonces $LR = 1$.

A través del punto crítico del LR se puede obtener un intervalo de apoyo de valores (*supported range*) del parámetro de estudio. Los valores del parámetro con LR por arriba del punto crítico se reportan como los valores de apoyo, mientras que los valores del parámetro con LR por debajo de este punto no son de apoyo (véase la Figura 1). La selección del punto crítico es por conveniencia. Un valor recomendado para el punto crítico es 0.258, por su correspondencia en ciertas condiciones con los intervalos de confianza al 90%. Asumiendo un estudio sobre una variable con distribución binomial bajo los siguientes resultados, $x = 4$, $n = 10$, donde el estimador de máxima verosimilitud de π es $4/10$, la amplitud de apoyo para el punto crítico de 0.258 es: 0.17, 0.65 (2). Esto indica que los valores de π con las razones de verosimilitud entre 0.17 y 0.65 tienen una probabilidad

de al menos un 25.8% con respecto a la probabilidad de los datos (función de verosimilitud) con π igual a 0.4. Contrario a los límites de confianza, uno puede interpretar los intervalos de verosimilitud sin hacer referencia a otros datos más allá de los datos observados (3).

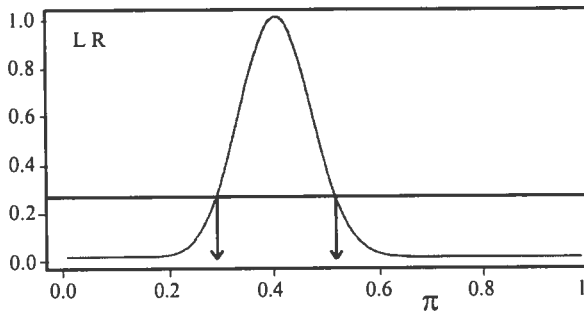


Figura 1. Razón de Verosimilitud para el caso Binomial

Estimación por el Método Bayesiano

Otra alternativa de estimación es mediante el uso de métodos bayesianos. Estos métodos utilizan la función de verosimilitud para actualizar una distribución a priori del parámetro y así, crear una distribución a posteriori (5,6).

La distribución a priori refleja el nivel de conocimiento o creencia (objetivo o subjetivo) que se tiene sobre el parámetro antes de observar los resultados muestrales. Esta distribución puede estar basada en información externa al estudio; por ejemplo, basada en opinión médica de expertos, revisión de literatura sobre estudios similares respecto al tema bajo consideración, o meta-análisis. Después de observados los datos, se utiliza la distribución resultante (distribución a posteriori) para determinar la probabilidad (objetiva o subjetiva) de que el parámetro está dentro de ciertos límites, a base de los datos observados y el conocimiento a priori del investigador sobre el parámetro o la medida de interés. En la estadística bayesiana, estos límites se identifican como límites de credibilidad (2). Debido a que la distribución a priori puede ser subjetiva, ya que puede variar de investigador a investigador, la distribución a posteriori puede resultar diferente. Por lo tanto, la objetividad del método bayesiano recae principalmente en la regla para modificar las probabilidades a priori a base de los datos observados (2). Sin embargo, cuando los resultados muestrales son convincentes, la influencia de la distribución a priori sobre la distribución a posteriori es relativamente pequeña. Por consiguiente, si dos investigadores tuvieran opiniones previas diferentes, antes del estudio, una evidencia muestral contundente los deberá llevar a una conclusión similar a posteriori.

Una de las ventajas de la estadística bayesiana es que provee respuestas probabilísticas a preguntas como ¿Está el verdadero parámetro entre 30% y 50%? Una posible respuesta bayesiana sería: a base de los datos observados y mi información previa sobre el parámetro, diría con un 90% de probabilidad que el parámetro está dentro de estos límites.

Como ilustración del teorema de Bayes, supongamos que π representa la proporción de pacientes que responde a un tratamiento, $0 < \pi < 1$. Es conveniente, desde el punto de vista estadístico, asumir que la distribución previa de π es:

$$p(\pi) \propto \pi^{a-1} (1 - \pi)^{b-1}.$$

Esta función es conocida como la distribución beta[a, b].

Luego de examinar n pacientes, se observa que x de ellos responden al tratamiento y $n - x$ no responden. La función de verosimilitud entonces, es:

$$p(\mathbf{x} | \pi) \propto \pi^x (1 - \pi)^{n-x}.$$

Por consiguiente, la distribución revisada es, también, beta[a+x, b+n-x]:

$$p(\pi | \mathbf{x}) \propto \pi^{a+x-1} (1 - \pi)^{b+n-x-1}.$$

Así, como en el caso ilustrado arriba, en que $x = 4$ y $n - x = 6$, asumimos una distribución uniforme para π en el intervalo 0 a 1; es decir, una distribución a priori beta [1, 1]. Luego de observados los 4 éxitos y los 6 fracasos, la distribución a posteriori de π será beta[5, 7] (véase la Figura 2); con media posterior $\mu(\pi | \mathbf{x}) = 5/12 = 0.42$ y desviación estándar posterior (5):

$$\sigma(\pi | \mathbf{x}) = \sqrt{\frac{5 \cdot 7}{12^2 \cdot 13}} = 0.1367$$

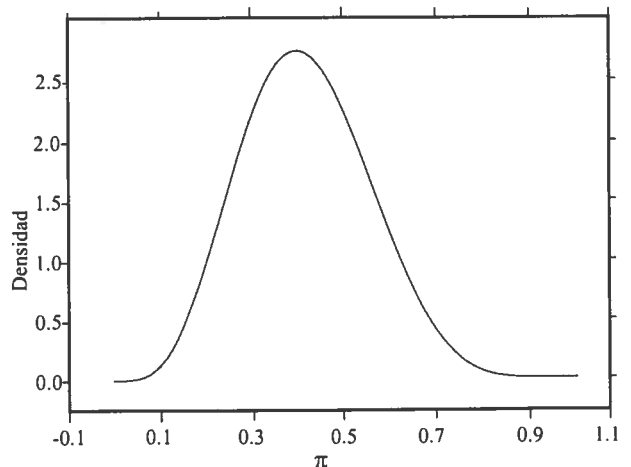


Figura 2. Distribución Beta [5, 7]

Podemos obtener los límites de credibilidad para π mediante integración numérica de la distribución posterior o mediante simulación generando muestras aleatorias de una distribución beta (5, 7). Por ejemplo, si utilizamos el programa estadístico *S-plus*, podemos generar, digamos, 1000 observaciones aleatorias de la distribución beta y de éstas, obtener las percentilas 5 y 95. Con ellas, obtenemos unos límites creíbles a 90% para π . A continuación ilustramos las instrucciones y resultados de este procedimiento, usando el programa *S-plus*:

```
> x <-rbeta(1000, 5, 7) ... [para generar aleatoriamente  
las 1000 observaciones de la distribución beta]  
> quantile(x, probs = .05) ... [obtener la percentila 5]  
5%  
0.1934692  
> quantile(x, probs = .95) ... [obtener la percentila 95]  
95%  
0.6493657
```

Así, que el intervalo creíble para π es:

$$pr(0.19 < \pi \mid \mathbf{x} < .65) = 0.90.$$

Otro argumento para utilizar los intervalos bayesianos es que éstos son más estrechos (más precisos) que los intervalos de confianza tradicionales con el mismo nivel de confianza (3). En general, es recomendable utilizar éstos cuando exista una clara justificación de su uso y, además, existan los programas de computación adecuados para llevar a cabo este análisis. El programa *BUGS* (*Bayesian Inference Using Gibbs Sampling*) está diseñado para realizar análisis en modelos estadísticos complejos. Este programa puede ser obtenido mediante *Internet* en la dirección siguiente: www.mrc-bsu.cam.ac.uk/bugs. Otro programa estadístico es *First Bayes*

(www.maths.nott.ac.uk/personal/ao/1b.html), el cual presenta una introducción a la enseñanza de la estadística bayesiana (7).

Conclusión

Con los avances en la tecnología moderna, la aplicación de los métodos de estimación que utilizan la función de verosimilitud y el paradigma bayesiano ha crecido y se ha expandido en diferentes campos. Sin embargo, sus aplicaciones en las publicaciones científicas en el campo de la epidemiología no son muy frecuentes. Entendemos que con el desarrollo de nuevos programas de computación y la difusión de estos métodos, sus aplicaciones irán en aumento, por lo que los investigadores de la salud tendrán nuevas perspectivas de análisis y evaluación de datos estadísticos.

Referencias

- 1) Collett D. Modeling binary data. London: Chapman and Hall; 1991.
- 2) Clayton D, Hills M. Statistical models in epidemiology. Oxford: Oxford University Press; 1993.
- 3) Rothman K, Greenland S. Modern epidemiology. Second Edition. Philadelphia: Lippincott-Raven Publishers; 1998.
- 4) Gordis L. Epidemiology. Philadelphia: WB Saunders Company; 1996.
- 5) Gelman A, Carlin J, Stern H, Rubin D. Bayesian data analysis. London: Chapman and Hall; 1995.
- 6) Sivia DS. Data analysis: a Bayesian tutorial. Oxford: Oxford University Press; 1997.
- 7) Guzmán J. Introducción a ensayos clínicos: una perspectiva bayesiana. Departamento de Bioestadística y Epidemiología, Escuela Graduada de Salud Pública, Recinto de Ciencias Médicas, Universidad de Puerto Rico. Abril, 2000.