

Inferencia estadística en investigaciones de salud pública a través de intervalos de confianza

André Chocó
Unidad de Investigación
Clínica de Enfermedades Infecciosas

INTRODUCCIÓN

Desde hace más de dos décadas, se ha cuestionado el uso de pruebas de hipótesis en el análisis de datos, atendiendo a los muchos puntos débiles que estos presentan y a la existencia de otras alternativas de análisis como lo son la estadística bayesiana y los métodos multivariados de análisis.

Los intervalos de confianza, a pesar que pertenecen a la misma tradición clásica de la estadística, aún gozan de suficiente respeto por la comunidad científica al ser más informativos que la sola presentación de un valor p de una prueba de hipótesis.

En artículo se hablará de los intervalos de confianza como una alternativa adecuada para la inferencia estadística en estudios de salud pública.

Importancia de los intervalos de confianza

En estudios médicos los investigadores deberían usualmente estar interesados en determinar el tamaño de la diferencia de una variable respuesta medida en ambos grupos, más que la simple indicación de que hay no diferencia estadísticamente significativa (1).

Los intervalos de confianza presentan un rango de valores, sobre la base de una muestra de datos, en los cuales el valor poblacional para una diferencia o un parámetro es probable que esté contenido (1–3).

En casos en los que el efecto estimado del tratamiento es pequeño, el intervalo de confianza indica cuando el beneficio de tratamiento clínicamente valioso sigue siendo plausible a la luz de los datos y puede ayudar a evitar confundir la falta de evidencia de efectividad con evidencia de falta de efectividad. Son especialmente utilizados en ensayos clínicos y meta análisis (1).

Los intervalos de confianza deberían utilizarse para presentar los resultados más importantes de una investigación, por tanto, deberían incluirse en el resumen y las conclusiones de una publicación científica (1,4).

Cómo se interpreta un intervalo de confianza

Nosotros estudiamos una muestra, al resumir los datos obtenemos estadísticos, o valores muestrales como la media muestral o la proporción muestral que nos dan una idea del correspondiente parámetro, es decir el valor de la media o la proporción calculado hipotéticamente en base al estudio de todos los individuos de la población (5–7).

Los intervalos de confianza indican la in(precisión) con que la muestra del estudio estima el valor del parámetro (1).

Otra forma de interpretar alude a determinar los resultados que tomaría la variable de interés si se repitiera el estudio en múltiples ocasiones. Lo correcto es indicar que, al calcular un intervalo de confianza, en el 95% de las veces el verdadero parámetro podría estar contenido en el rango de valores del intervalo calculado; y por tanto en un 5% de las veces podría no estar en ese rango de valores (1,8).

Un ejemplo

Si: 34 a 49 años es el intervalo de confianza del 95% de la media poblacional de la edad de pacientes con vih que asisten a una clínica, para una muestra de tamaño “ n ”, observamos que el verdadero valor de la edad de todos los pacientes con vih que reciben atención médica, estimado a través de una muestra de tamaño “ n ”,

podría encontrarse en 95 de 100 veces entre 34 a 49 años, eso implica un rango de posibilidades discretas de 16 edades diferentes.

En un 5% de las veces se espera que el verdadero valor no esté incluido en este intervalo. Basados en los datos muestrales, tenemos una confianza del 95% de que el verdadero parámetro correspondiente a la edad promedio esté contenido en el intervalo de 34 a 49 años.

Intervalos de confianza y tamaños de muestra

A medida que aumentamos el tamaño de muestra, podemos disponer de mayor evidencia y por tanto realizar una estimación con mayor validez, esto ocurre porque el intervalo de confianza depende en parte, del error estándar, el cual se obtiene al dividir la desviación estándar muestral entre la raíz del tamaño de la muestra, es decir una vez la raíz del tamaño muestral aumenta en relación a la desviación estándar, el ancho del intervalo será menor. El otro aspecto que influye sobre la amplitud del intervalo es el nivel de confianza, a mayor nivel de confianza, si se mantiene inalterado el error estándar, la amplitud del intervalo aumenta (1).

Podemos decir entonces que, si dejamos constante el nivel de confianza y la desviación estándar observada, a medida que aumenta el tamaño de muestra la amplitud de un intervalo de confianza se reduce, y por tanto, la precisión de la estimación aumenta.

La potencia aumenta también cuando se aumenta el tamaño del efecto, ya que disminuye la probabilidad de error beta (aceptar una hipótesis nula falsa), sin embargo, podría aumentar el riesgo de cometer un error alfa (rechazar una hipótesis nula cierta). Hay que recordar que la amplitud de un intervalo decrece a una tasa equivalente de raíz del tamaño de muestra (6,9).

Por tanto, un intervalo de confianza muy amplio podría sugerirnos que la muestra utilizada no fue la adecuada en tamaño y un intervalo de confianza muy estrecho, que utilizamos demasiados pacientes, y esto, en ciencias de la salud podría resultar no ético, por la posibilidad de exponer a más de los individuos necesarios a tratamientos con potenciales riesgos colaterales (1,6,7,10).

A continuación, se muestra la forma de calcular el intervalo de confianza de un estimador puntual, donde SE, es el error estándar, o sea, como se mencionó anteriormente, en el caso de estimación de una media, el cociente de la desviación estándar muestral entre la raíz cuadrada del tamaño de muestra (1):

$$\text{estimador puntual} - (z_{\alpha/2} \times SE) \text{ a } \text{estimador puntual} + (z_{\alpha/2} \times SE)$$

Tipos de intervalos de confianza

Si bien, los intervalos de confianza pueden pensarse como métodos complementarios de las pruebas hipótesis, ya que aportan información adicional; sin embargo, algunos Estadísticos y Epidemiólogos prefieren sustituir el uso de pruebas de hipótesis por intervalos de confianza, con el argumento que estos últimos pueden aplicarse a la solución de cualquier problema que resuelven las pruebas de hipótesis (8,9).

Ahora se expondrá la clasificación de los intervalos de confianza atendiendo al tipo de problema que pretenden resolver.

A continuación, se enumeran algunos de los muchos intervalos de confianza existentes y se anota en paréntesis la prueba estadística que estos sustituyen o complementan:

A. Estimación de un parámetro:

- De una media poblacional (Prueba de T Student de una media)
- De una proporción poblacional (Prueba de Z para una muestra, test binomial)
- De una mediana poblacional (podría ser el equivalente o el complemento de una prueba de rangos signados de Wilcoxon para una mediana, o la prueba de signos para una muestra)
- De la sensibilidad y la especificidad de una prueba diagnóstica (1,3).

B. Estimación de una diferencia poblacional:

- De medias poblacionales, dos muestras independientes (T de Student)
- De medias poblacionales, dos muestras relacionadas (T de Student pareada)
- De proporciones poblacionales, dos muestras independientes (Prueba de Z de proporciones independientes, test de ji cuadrado, prueba exacta de Fisher)

- De proporciones poblacionales, dos muestras relacionadas (Prueba de Z de proporciones relacionadas, test de McNemar)
- De medianas poblacionales, dos muestras independientes (Prueba de Mann-Whitney)
- De medianas poblacionales, dos muestras relacionadas (Prueba de rangos signados de Wilcoxon)
- De la diferencia de dos coeficientes de correlación (Transformación de Fisher previo a la aplicación de un test de Z)
- De la diferencia de dos curvas roc (Covarianza de dos áreas bajo la curva correlacionadas)
- De la diferencia de pendientes poblacionales de dos curvas de regresión
- De las diferencias de medias en comparaciones múltiples (ANDEVA + métodos de comparaciones múltiples) (1-3,11).

C. Estimación del tamaño del efecto

- De un riesgo relativo, de una OR, de una razón de prevalencias (Prueba de ji cuadrado, prueba exacta de Fisher)
- De una OR en un estudio de casos y controles pareados (Test de Mc Nemar)
- De un coeficiente de correlación poblacional (Transformación de Fisher y luego estadístico Z, test de T)
- De un Hazard Ratio en una regresión de Cox
- De una OR ajustada en una regresión logística (Estadística de Wald)
- De un coeficiente Kappa para concordancia entre dos variables categóricas (Test de T para coeficiente Kappa poblacional)
- De la estimación del área bajo una curva roc (Prueba binomial, Test U de Mann-Whitney para una muestra)
- De la pendiente poblacional de una regresión (Prueba de T para una pendiente poblacional) (1,3,11).

D. Otros

- De la predicción de un valor medio de Y en función de un valor de X (Prueba de T de Student para una muestra aplicada a los datos de Y) (1).

BIBLIOGRAFÍA

1. Altman D, Machin D, Bryant T, Gardner M. *Statistics with Confidence*. 2nd ed. Reino Unido: British Medical Journal; 2000. 240 p.
2. Peacock J, Peacock P. *Oxford Handbook of Medical Statistics*. Reino Unido: Oxford University Press; 2011. 540 p.

3. Blair C, Taylor R. *Bioestadística*. México: Pearson Educación, S.A.; 2008. 552 p.
4. Du Prel J-B, Hommel G, Röhrig B, Blettner M. Confidence Interval of P-Value. [cited 2017 Apr 25]; Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2689604/pdf/Dtsch_Arztebl_Int-106-0335.pdf
5. Silva L. *Diseño razonado de muestras y captación de datos para investigación sanitaria*. Madrid: Díaz de Santos; 2000. 332 p.
6. Velasco Rodríguez V, Martínez Ordaz V, Roiz Hernández J, Huazano García F, Nieves Rentería A. *Muestreo y tamaño de muestra: Una guía práctica para personal de salud que realiza investigación*. Buenos Aires: e-libro.net; 2002. 172 p.
7. Lwanga S, Lemeshow S. *Determinación del tamaño de las muestras en los estudios sanitarios: Manual práctico*. Ginebra: Organización Mundial de la Salud; 1991. 80 p.
8. Scotto MG, Tobías-Garcés A, Stat C. Interpretando correctamente en salud pública estimaciones puntuales, intervalos de confianza y contrastes de hipótesis. 2003 [cited 2017 Apr 24];45(6). Available from: <http://>
9. Argimon Pallas J, Jiménez Villa J. *Métodos de investigación clínica y epidemiológica*. 3rd ed. Madrid: Elsevier; 2000. 393 p.
10. Ruiz A, Morillo L. *Epidemiología Clínica: Investigación clínica aplicada*. Bogotá: Editorial Médica Panamericana; 2004. 576 p.
11. Glantz S. *Bioestadística*. 6th ed. México, D.F.: Mc Graw Hill; 2006.