

Epidemiología Clínica

Epidemiología Clínica: Revisando la literatura médica. Significancia estadística: $p < 0,05$

JUAN PABLO CHALCO ORREGO*

En el proceso de revisión de un artículo original nos encontramos frecuentemente con dos tipos de preguntas: ¿Cuánto sabemos del tema del artículo? y ¿Cómo se valora la forma en que se presenta o diseña un artículo o trabajo de investigación?. Sobre la primera pregunta es obvio que tenemos que recurrir a nuestros conocimientos anteriores o experiencia y revisar un poco sobre el tema en términos generales (acudiendo a revisiones formales serias o preguntando a expertos en el tema); pero la segunda interrogante requiere de un poco más que preguntas simples. Necesitamos de conocimientos de estadística y epidemiología clínica que trataremos de presentarles abreviadamente en el transcurso de los siguientes números.

Es cierto que actualmente existe más de una revisión del tema por lo que nos enfocaremos a conceptos importantes y a la vez poco entendidos. Para mayor información se les sugiere revisar las referencias al final del artículo.

El concepto más ampliamente usado y a la vez menos entendido es el p y su valoración estadística. La gran mayoría de los artículos originales publicados presentan en sus resultados valores de p calculados de diversas maneras en un afán de encontrar “significancia estadística”. Pero empezamos por el principio.

En medicina es frecuente el querer comparar resultados de dos o más grupos ya sea en

sus características basales o en el resultado de una o varias intervenciones. Para tal fin se empieza primero seleccionando adecuadamente una muestra representativa de la población. La **teoría del muestreo** permite hacer inferencias a la población de origen a partir de estudios realizados en una o más muestras extraídas de esta a pesar del riesgo del error o imprecisión ya sean debidas **al azar o la variabilidad biológica del fenómeno a estudiar**. A mayor tamaño muestral, mayor es la precisión y la variabilidad explicada por el azar decrece. Esta posibilidad de error o falta de precisión se evita aumentando el tamaño de la muestra. La **posibilidad del azar** debe ser siempre buscada y evaluada al comparar grupos realizando test de hipótesis para probar la significancia estadística de la diferencia encontrada o construyendo intervalos de confianza para conocer la precisión de nuestra estimación dentro de una seguridad previamente definida. **El p es la posibilidad que un resultado (diferencias encontradas entre uno o más grupos) sea debida al azar; a menor valor ($p < 0,05$) más significancia estadística.**

Así empezamos nuestra prueba de significancia estadística comparando dos grupos con la **formulación de la hipótesis nula (H_0)**. Esta formula la igualdad de los resultados de dos muestras 1 y 2, es decir la no diferencia estadística entre dos resultados (medias o proporciones) y se escribe así:

$$x_1 = x_2 \text{ para medias o } p_1 = p_2 \text{ para proporciones}$$

* Médico Pediatra Asistente Medicina C - IESN.

Esta es la opción que generalmente queremos demostrar que es **falsa** cuando estamos probando una intervención en dos grupos comparables “...*existe diferencia significativa entre tratamiento A y tratamiento B*” o a veces la opción que queremos demostrar que es **verdadera** al comparar características basales de dos grupos previo a un ensayo “...*no existen diferencias significativas en las características basales del Grupo A y grupo Grupo B por lo que son comparables*”.

La **hipótesis alternativa (H_1)** propone a su vez que sí existe diferencias entre dos grupos de muestras (1 y 2) y se representa como:

$$x_1 \neq x_2 \text{ para medias o } p_1 \neq p_2 \text{ para proporciones}$$

Y recíprocamente, esta hipótesis es la opción que generalmente queremos demostrar es **verdadera** al evaluar una intervención en dos grupos comparables o que es **falsa** al comparar características basales de dos grupos antes de una intervención.

El segundo paso es **escoger el test adecuado para calcular el p** , generalmente prueba t de student para variables paramétricas (medias) y Chi cuadrado para variables no paramétricas (proporciones) por citar las más conocidas. Este paso es fundamental para la obtención adecuada del p . El p calculado de esta manera nos va a llevar a rechazar o aceptar la hipótesis nula. Acá podemos cometer varios errores y correr riesgos relacionados a ellos (Tabla 1).

Tabla 1.- Errores y riesgos al probar una hipótesis.

Error	Definición del Tipo de error	Riesgo asociado
Tipo I	Rechazar equivocadamente la hipótesis nula cuando es cierta	La probabilidad de un error tipo I = ?
Tipo II	Aceptar equivocadamente la hipótesis nula cuando es falsa	La probabilidad de un error tipo II = ?

El riesgo alfa es entonces la probabilidad de rechazar equivocadamente una hipótesis nula cuando es cierta, es decir, aceptar que existen diferencias significativas cuando no existen o que son debidas al azar. Por lo tanto, nos conviene minimizar este riesgo por lo que el tercer paso es **escoger el nivel de significancia o alfa** que no es sino el nivel máximo de cometer un error tipo I o alfa. Generalmente se escoge $p < 0,05$ que es menos del 5% la probabilidad de cometer este error o menos de 1 en 20 esta posibilidad. Cuando el nivel escogido de p es $< 0,01$ equivale a una probabilidad menor del 1% o 1 en 100. Un $p = 0,05$ significa una posibilidad de sólo 1 en 20 que los resultados obtenidos se deban al azar lo que parecería significativo. Viéndolo de otra manera tengo 19 de 20 posibilidades que la diferencia encontrada no se deba al azar, pero un $p = 0,06$ si bien no parecería significativo sigue siendo alta su probabilidad (17 de 18) de no deberse al azar, por lo que debería concluirse con este p la necesidad de un trabajo de investigación posterior. Por esto mismo usualmente no se escoge $p < 0,1$ pues esta probabilidad de cometer un error tipo I si es alta (10% ó 1 en 10). Hasta este momento hemos escogido el nivel de significancia, el test apropiado y formulamos la hipótesis nula. Esto se resume en la Figura 1.

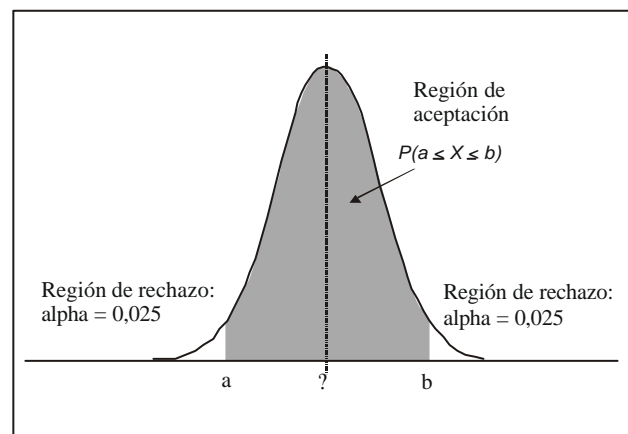


Figura 1.- Curva de distribución normal para una variable paramétrica (una media por ejemplo) con prueba de 2 colas (2,5% ó 0,025 cada lado).

La hipótesis alternativa establece $x_1 \neq x_2$ lo que dice en realidad que la diferencia encontrada puede ser mayor o menor. Se establecen entonces zonas de aceptación y rechazo de la hipótesis nula. La zona de rechazo es en realidad el nivel significativo del p ($<0,05$). El valor calculado del p puede caer en una de estas zonas. En este ejemplo implica una prueba de dos colas o dos zonas de rechazo de la hipótesis nula (p significativo), una cola para la posibilidad que el resultado sea menor hacia la izquierda y otra cola hacia la derecha cuando la diferencia encontrada es mayor. Como el nivel de significancia fue establecido en $p < 0,05$ este se distribuye simétricamente en las dos colas, cada una con 2,5% de área. De esta figura podemos inferir dos puntos importantes. Primero la zona de rechazo de la hipótesis nula varía con el nivel de significancia ($p < 0,05$ ó 5% del área de la curva es diferente a $p < 0,01$ ó 1% del área) siendo más preciso a menor área. En segundo lugar esta área varía también con la formulación de la hipótesis, ya que si se formula que la diferencia es mayor uno del otro: $x_1 > x_2$ ese 5% del área se va a un solo lado de la curva, lo que hace una prueba de una sola cola, variando la posibilidad de rechazo de la hipótesis nula cuando se calcule el p (Figura 2).

Por lo cual el cálculo del p y su nivel de significancia varía al escoger las hipótesis, el nivel de significancia y al elegir el test para

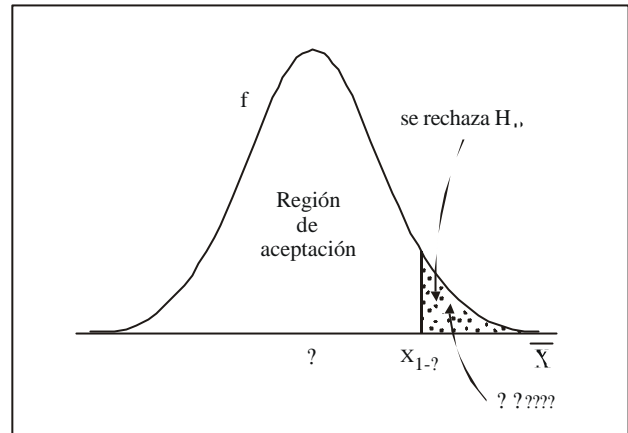


Figura 2.- Curva de distribución normal para una variable paramétrica con prueba una cola (5% ó 0,05).

calcularlo; esto último veremos en el siguiente número.

BIBLIOGRAFÍA

1. Sacket DL, Haynes RB, Guyat GH. Clinical Epidemiology: a basic science for Clinical Medicine. Little Brown:1991.
2. D.L. Sacket DL, Richardson WS, Rosenberg W. Evidence Based Medicine: How to Practice and Teach EBM. Churchill Livingstone: 1997.
3. Mould R. Introductory Medical Statics. Publishing Bristol & Philadelphia:1998.
4. Petrie A, Sabin C. Medical statics at a glance. Blackwell Science Ltd. 2000.