

Breve repaso de temas de estadística aplicada a la investigación médica

Cuauhtémoc Acoltzin Vidal*

ASOCIACIÓN ENTRE DOS CARACTERÍSTICAS NOMINALES

Cuando la variable dependiente (VD) y la variable independiente (VI, una o varias) son nominales se podrá identificar asociación entre ellas, sea en sentido de exposición o de protección.

En tal caso, a la variable de interés (VD) se le denomina efecto y a la VI se le dice factor de riesgo (FR) o de exposición. Es cierto que, en los ensayos clínicos, por lo general la VI ofrece protección (como un procedimiento terapéutico, por ejemplo) a pesar de lo cual se respeta el término acuñado de FR.

Para identificar tal asociación se recurre al cálculo de riesgo relativo (RR) (también conocido como razón del riesgo) o la razón de momios (OR, por *odds ratio*; llamada también razón dispar o razón de productos cruzados). Los momios son las proporciones en las cuales uno de los participantes recibe algo más sobre lo que le tocaría, por eso se dice también de disparidad.

El RR está indicado en estudios prospectivos: de cohorte o ensayo clínico. Es una razón de frecuencias relativas de: incidencia en personas expuestas (E), entre incidencia en personas no expuestas (E). Se hace partiendo de una tabla de distribución de 2 x 2 cuyas celdas se denominan A, B, C y D. Se divide la proporción de personas expuestas al riesgo, que tienen el efecto, entre la de personas no expuestas, que tienen el efecto:

$$RR = \frac{A/(A+B)}{C/(C+D)}$$

$$RR = 0.21$$

Ejemplo supuesto	EVC presente	EVC ausente	Sub totales	Riesgo real	X 100 = %
HTA en control	Celda A 294	Celda B 18,506	Celda n1 18,800	A/n1 0.0156	Proporción 1.56%
HTA sin control	Celda C 1,416	Celda D 17,384	Celda n2 18,800	C/n2 0.0753	Proporción 7.53%

HTA: hipertensión arterial; EVC: evento cerebral vascular.

La OR es una medida del tamaño de un efecto y puede ser reportada en estudios retrospectivos: de casos y controles, o prospectivos incluyendo metaanálisis. Se le dice de productos cruzados debido a que puede definirse como la relación del producto de las diagonales de la tabla de distribución:

$$OR = \frac{\frac{A}{A+C} / \frac{C}{A+C}}{\frac{B}{B+D} / \frac{D}{B+D}}$$

$$OR = \frac{A/C}{B/D}$$

$$OR = \frac{A/D}{B/C}$$

$$OR = 0.20$$

Se trata de una razón de frecuencias relativas de E con FR, entre frecuencias relativas de E sin FR. O número de casos en que la relación entre FR y E es real dividido entre el número de aquellos entre quienes no existe. Oakley y cols. explican como la OR de seis significa que seis personas experimentarán el efecto por cada uno que no (seis de cada siete), en

* Médico Cirujano, Cardiólogo y Maestro en Ciencias Médicas. Profesor de Estadística aplicada a la investigación en medicina, Maestría en Ciencias Médicas, Facultad de Medicina, Universidad de Colima.

cambio la OR de 0.20 significa un efecto de cada cinco que no (uno de seis).

Tanto el RR como la OR se complementan con cálculo de intervalo de confianza, por lo general de 95% (IC de 95%) para extrapolar a esa proporción de la población de referencia.

Para interpretar el RR se observa si la cifra puntual es menor a uno, en cuyo caso se dice que protege; o si es mayor a uno en cuyo caso se dice que expone. Entre más alejado de uno mayor será la asociación. También resulta importante observar el IC de 95%, porque si sobrepasa la línea de la gráfica que señala el uno restará vigencia a la supuesta observación aunque el valor p sea significativo, ya que éste sólo indica la probabilidad de intersección de dos subconjuntos de una población. También importa que el intervalo de confianza sea corto porque esto contribuye a validar la prueba. Con los datos del ejemplo supuesto IC de 95% en caso de RR es: de 0.17 a 0.22; en caso de OR es de 0.18 a 0.24. Nótese que ni siquiera se acercan a uno, por lo tanto se concluye que el control de hipertensión arterial (HTA) protege de evento vascular cerebral (EVC) de manera confiable.

Para definir la magnitud de los cambios ocurridos en los estudios de intervención sancionados por la modificación de riesgo y la posibilidad de aplicar los resultados en el mundo real se recurre a tres pruebas adicionales: reducción absoluta de riesgo (RAR), reducción relativa de riesgo (RRR) y número a tratar (NT).

La RAR representa la proporción en que se ha modificado la presentación de un evento en función del FR; resulta de restar el riesgo observado del riesgo esperado. En nuestro ejemplo la reducción absoluta de riesgo es 5.97%.

La RRR significa el peso que tal modificación tiene en el resultado; se calcula dividiendo RAR entre el riesgo esperado, que en el caso supuesto es 79.28%. Es decir, que 5.97 es el 72.28% de 7.53.

El NT significa el número mínimo de personas quienes se deben exponer al FR para que ocurra una vez el efecto; es la recíproca de la RAR, es decir $1/\text{RAR} = 1/0.0597 = 16.75$. Significa que será necesario controlar bien la HTA de 16 o 17 personas para evitar un caso de EVC.

La prueba de significancia estadística apropiada para el análisis de asociación entre variables nominales es la llamada χ^2 (ji cuadrada –o chi cuadrada–), que tiene varias aplicaciones. La χ^2 de Pearson (para muestras no relacionadas) se basa en la

suma de comparación entre los datos esperados y los obtenidos en cada celda de la tabla de 2 x 2 y el resultado se compara con tablas elaboradas de antemano concediendo tantos grados de libertad cuantas columnas tenga la tabla, menos una multiplicada por el número de renglones de la tabla menos uno. Para el ejemplo supuesto el valor $p = 0.0000001$.

Si la muestra es pequeña o el valor esperado de alguna casilla de la tabla de distribución es menor a cinco, se preferirá la variedad descrita por Fisher y Yates denominada «prueba exacta».

Las dos pruebas se pueden hacer con cualquier paquete de análisis estadístico, porque dada su relevancia son cada vez más utilizadas. La OPS ofrece su paquete Epi Info 7 que es libre y se puede obtener en: <http://www.cdc.gov/EpiInfo/> (en inglés) o si se prefiere en español: <http://huespedes.cica.es/huespedes/epiinfo/>. Este paquete facilita la realización de estas pruebas a partir del comando STATCALC que ofrece la tabla de 2 x 2 y conforme se van escribiendo los datos en las celdas va mostrando resultados de RR, OR, χ^2 (en varias modalidades) y valor p .

La prueba χ^2 para proporciones con más de dos grupos de observación resulta muy útil, pero los paquetes de estadísticas no la hacen. El procedimiento es laborioso pero la fórmula se puede transformar en algoritmo con Excel de Microsoft Office. Hela aquí:

$$\frac{1}{P \cdot Q} \sum Ni(p - p')^2$$

[P = se usa aquí en lugar de «p testada» que es proporción media de que ocurra un suceso y se obtiene sumando los porcentajes divididos entre 100; Q = se usa aquí en lugar de «q testada», proporción media de que no ocurra el suceso se obtiene igual para expresar como fracción; Σ = sumar; Ni = tamaño de la muestra de cada subgrupo; p = proporción observada del suceso expresada como fracción].

Si los grupos que se analizan están relacionados (antes-después, pareados o emparejados) se usan variedades distintas de la prueba: Para dos grupos (diseño frecuente) es la de McNemar, muy fácil de hacer: $[(B-C)-1]^2/2 + C$. (Recordar los nombres de las celdas de 2 x 2).

Para más de dos grupos Cochran diseña la prueba Q. También se hace directamente con Epi Info.

Por último, cuando hay factores de confusión o se quieren analizar cambios ocurridos en lapsos de

tiempo, como la sobrevida, se recurre a la prueba de Mantel y Heanszel, que también se puede hacer con paquetes de análisis estadístico.

BIBLIOGRAFÍA COMPLEMENTARIA

1. Oakley-Davies HT, Kinloch-Crombie I, Tavakoli M. ¿Cuándo puede llevar la razón de momios a conclusiones erróneas? BMJ Edición Latinoamericana. 1998; 6: 167-170.
2. Cook RJ, Sacket DT. El número necesario para tratar; una medida clínicamente útil. BMJ Latinoamericana. 1995; 3: 227-229.
3. Dawson-Saunders B, Trapp RG. Bioestadística médica. México: Editorial El Manual Moderno S.A. de C.V.; 1993. (capítulos 9, 11 y 12).
4. Riegelman RK, Hirsch RP. Cómo estudiar un estudio y probar una prueba: lectura crítica de la literatura médica. Publicación científica 531. Washington: OPS; 1995.
5. Castilla-Serna L, Cravito J. Estadística simplificada para la investigación en ciencias de la salud. México: Trillas; 1991.

Dirección para correspondencia:

Cuauhtémoc Acoltzin Vidal

E-mail: cuauhtemoc_acoltzin@ucol.mx