

Asociación de variables cualitativas: test de Chi-cuadrado

Autores:

Salvador Pita Fernández⁽¹⁾, Sonia Pértega Díaz⁽²⁾

(1) Médico de Familia. Centro de Salud de Cambre (A Coruña).

(2) Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario Juan Canalejo (A Coruña).

Actualizada el 03/11/2004.

En la investigación biomédica nos encontramos con frecuencia con datos o variables de tipo cualitativo, mediante las cuales un grupo de individuos se clasifican en dos o más categorías mutuamente excluyentes. Las proporciones son una forma habitual de expresar frecuencias cuando la variable objeto de estudio tiene dos posibles respuestas, como presentar o no un evento de interés (enfermedad, muerte, curación, etc.). Cuando lo que se pretende es comparar dos o más grupos de sujetos con respecto a una variable categórica, los resultados se suelen presentar a modo de tablas de doble entrada que reciben el nombre de tablas de contingencia. Así, la situación más simple de comparación entre dos variables cualitativas es aquella en la que ambas tienen sólo dos posibles opciones de respuesta (es decir, variables dicotómicas). En esta situación la tabla de contingencia se reduce a una tabla dos por dos como la que se muestra en la [Tabla 1](#)^{1,2}.

Supongamos que se quiere estudiar la posible asociación entre el hecho de que una gestante fume durante el embarazo y que el niño presente bajo peso al nacer. Por lo tanto, se trata de ver si la probabilidad de tener bajo peso es diferente en gestantes que fumen o en gestantes que no fumen durante la gestación. Para responder a esta pregunta se realiza un estudio de seguimiento sobre una cohorte de 2000 gestantes, a las que se interroga sobre su hábito tabáquico durante la gestación y se determina además el peso del recién nacido. Los resultados de este estudio se muestran en la [Tabla 2](#).

En la [Tabla 1](#), a, b, c y d son las frecuencias observadas del suceso en la realidad de nuestro ejemplo de estudio (43, 207, 105 y 1647), siendo n (2000) el número total de casos estudiados, y a+b, c+d, a+c y b+d los totales marginales. En el ejemplo, a+b=250 sería el número total de mujeres fumadoras durante el embarazo, c+d=1750 el número total de mujeres no fumadoras, a+c=148 el número de niños con bajo peso al nacer y b+d=1852 el número de niños con peso normal al nacimiento.

Ante una tabla de contingencia como la anterior pueden plantearse distintas cuestiones. En primer lugar, se querrá determinar si existe una relación estadísticamente significativa entre las variables estudiadas. En segundo lugar, nos interesará cuantificar dicha relación y estudiar su relevancia clínica. Esta última cuestión podrá resolverse mediante las denominadas medidas de asociación o de efecto (riesgo relativo (RR), odds ratio (OR), reducción absoluta del riesgo (RAR)), que ya han sido abordadas en otros trabajos^{3,4}. Por otro lado, para responder a la primera pregunta, la metodología de análisis de las tablas de contingencia dependerá de varios aspectos como son: el número de categorías de las variables a comparar, del hecho de que las categorías estén ordenadas o no, del número de grupos independientes de sujetos que se estén considerando o de la pregunta a la que se desea responder⁵.

Existen diferentes procedimientos estadísticos para el análisis de las tablas de contingencia como la prueba χ^2 , la prueba exacta de Fisher, la prueba de McNemar o la prueba Q de Cochran, entre otras.

En este artículo se expondrá el cálculo e interpretación de la prueba χ^2 como método estándar de análisis en el caso de grupos independientes^{1,2,5,6}.

La prueba χ^2 en el contraste de independencia de variables aleatorias cualitativas.

La prueba χ^2 permite determinar si dos variables cualitativas están o no asociadas. Si al final del estudio concluimos que las variables no están relacionadas podremos decir con un determinado nivel de confianza, previamente fijado, que ambas son independientes.

Para su cómputo es necesario calcular las frecuencias esperadas (aquellas que deberían haberse observado si la hipótesis de independencia fuese cierta), y compararlas con las frecuencias observadas en la realidad. De modo general, para una tabla r x k (r filas y k columnas), se calcula el valor del

estadístico χ^2 como sigue:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (1)$$

donde:

- O_{ij} denota a las frecuencias observadas. Es el número de casos observados clasificados en la fila i de la columna j .
- E_{ij} denota a las frecuencias esperadas o teóricas. Es el número de casos esperados correspondientes a cada fila y columna. Se puede definir como aquella frecuencia que se observaría si ambas variables fuesen independientes.

Así, el estadístico χ^2 mide la diferencia entre el valor que debiera resultar si las dos variables fuesen independientes y el que se ha observado en la realidad. Cuanto mayor sea esa diferencia (y, por lo tanto, el valor del estadístico), mayor será la relación entre ambas variables. El hecho de que las diferencias entre los valores observados y esperados estén elevadas al cuadrado en (1) convierte

cualquier diferencia en positiva. El test χ^2 es así un test no dirigido (test de planteamiento bilateral), que nos indica si existe o no relación entre dos factores pero no en qué sentido se produce tal asociación.

Para obtener los valores esperados E_{ij} , estos se calculan a través del producto de los totales marginales dividido por el número total de casos (n). Para el caso más sencillo de una tabla 2x2 como la [Tabla 1](#), se tiene que:

$$E_{11} = \frac{(a+b) \times (a+c)}{n} \quad E_{21} = \frac{(c+d) \times (a+c)}{n}$$

$$E_{12} = \frac{(a+b) \times (b+d)}{n} \quad E_{22} = \frac{(c+d) \times (b+d)}{n}$$

Para los datos del ejemplo en la Tabla 2 los valores esperados se calcularían como sigue:

$$E_{11} = \frac{148 \times 250}{2000} = 18,5 \quad E_{21} = \frac{148 \times 1750}{2000} = 129,5$$

$$E_{12} = \frac{1852 \times 250}{2000} = 231,5 \quad E_{22} = \frac{1852 \times 1750}{2000} = 1620,5$$

De modo que los valores observados y esperados para los datos del ejemplo planteado se muestran en la [Tabla 3](#).

El valor del estadístico χ^2 , para este ejemplo en concreto, vendría dado entonces como:

$$\chi^2 = \frac{(43-18,5)^2}{18,5} + \frac{(207-231,5)^2}{231,5} + \frac{(105-129,5)^2}{129,5} + \frac{(1645-1620,5)^2}{1620,5} = 40,04$$

A la vista de este resultado, lo que tenemos que hacer ahora es plantear un contraste de hipótesis entre la hipótesis nula:

H_0 : No hay asociación entre las variables (en el ejemplo, el bajo peso del niño y el hecho de fumar durante la gestación son independientes, no están asociados).

Y la hipótesis alternativa:

H_a : Sí hay asociación entre las variables, es decir, el bajo peso y el fumar durante la gestación están asociados.

Bajo la hipótesis nula de independencia, se sabe que los valores del estadístico χ^2 se distribuyen según una distribución conocida denominada ji-cuadrado, que depende de un parámetro llamado "grados de libertad" (g.l.). Para el caso de una tabla de contingencia de r filas y k columnas, los g.l. son igual al producto del número de filas menos 1 (r-1) por el número de columnas menos 1 (k-1). Así, para el caso en el que se estudie la relación entre dos variables dicotómicas (Tabla 2x2) los g.l. son 1.

De ser cierta la hipótesis nula, el valor obtenido debería estar dentro del rango de mayor probabilidad según la distribución ji-cuadrado correspondiente. El valor-p que usualmente reportan la mayoría de paquetes estadísticos no es más que la probabilidad de obtener, según esa distribución, un dato más extremo que el que proporciona el test o, equivalentemente, la probabilidad de obtener los datos observados si fuese cierta la hipótesis de independencia. Si el valor-p es muy pequeño (usualmente se considera $p < 0.05$) es poco probable que se cumpla la hipótesis nula y se debería de rechazar.

En la [Tabla 4](#), se determinan los grados de libertad (en la primera columna) y el valor de α (en la primera fila). El número que determina su intersección es el valor crítico correspondiente. De este modo,

si el estadístico χ^2 que se obtiene toma un valor mayor se dirá que la diferencia es significativa. Así, para una seguridad del 95% ($\alpha = 0.05$) el valor teórico de una distribución ji-cuadrado con un grado de libertad es 3,84. Para $\alpha = 0.01$ es de 6,63 y para $\alpha = 0.005$ es de 7,88. Como quiera que en el cálculo del χ^2 en el ejemplo obtuvimos un valor de 40,04, que supera al valor para $\alpha = 0.005$, podremos concluir que las dos variables no son independientes, sino que están asociadas ($p < 0.005$). Por lo tanto, a la vista de los resultados, rechazamos la hipótesis nula (H_0) y aceptamos la hipótesis alternativa (H_a) como probablemente cierta.

Para el caso de una Tabla 2x2, la expresión (1) del estadístico χ^2 puede simplificarse y obtenerse como:

$$\chi^2 = \frac{n(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}$$

Cuando el tamaño muestral es reducido la utilización de la distribución ji-cuadrado para aproximar las frecuencias puede introducir algún sesgo en los cálculos, de modo que el valor del estadístico χ^2 tiende a ser mayor. En ocasiones se utiliza una corrección para eliminar este sesgo que, para el caso de tablas 2x2 se conoce como la corrección de Yates:

$$\chi_Y^2 = \frac{n \left(|ad - bc| - \frac{n}{2} \right)^2}{(a + b)(c + d)(a + c)(b + d)}$$

En el ejemplo previo el cálculo del estadístico χ^2 con la corrección de Yates nos daría un valor de $\chi_Y^2 = 38,43$ ($p < 0.01$) en lugar de $\chi^2 = 40,04$. No existe consenso en la literatura sobre la utilización o no de esta corrección conservadora de Yates, que con muestras reducidas dificulta rechazar la hipótesis nula, si bien el efecto es prácticamente imperceptible cuando se trabaja con muestras de mayor tamaño.

No obstante, conviene mencionar que la utilización de la corrección de Yates no exime de ciertos requerimientos acerca del tamaño muestral necesario para la utilización del estadístico χ^2 . Como norma general, se exigirá que el 80% de las celdas en una tabla de contingencia deban tener valores esperados mayores de 5. Así, en una tabla 2x2 será necesario que todas las celdas verifiquen esta condición, si bien en la práctica suele permitirse que una de ellas muestre frecuencias esperadas ligeramente por debajo de este valor. En aquellos casos en los que no se verifique este requisito existe un test, propuesto por R.A. Fisher, que puede utilizarse como alternativa al test χ^2 y que se conoce

como test exacto de Fisher. El procedimiento consiste en evaluar la probabilidad asociada a todas las tablas 2x2 que se pueden formar con los mismos totales marginales que los datos observados, bajo el supuesto de independencia. Los cálculos, aunque elementales, resultan algo engorrosos, por lo que no se incluirán en este trabajo, siendo múltiples las referencias que se pueden consultar a este respecto^{5,6}.

Para finalizar, recalcar que existen otros métodos estadísticos que nos permiten analizar la relación entre variables cualitativas, y que vienen a complementar la información obtenida por el estadístico χ^2 . Por una parte, el análisis de los residuos estandarizados permitirá constatar la dirección en que se da la relación entre las variables estudiadas. A su vez, existen también otras medidas de asociación, muchas de las cuales resultan especialmente útiles cuando alguna de las variables se mide en una escala nominal u ordinal, que permiten cuantificar el grado de relación que existe entre ambos factores⁷.

Bibliografía

1. Fleiss J L. Statistical Methods for rates and proportions. 3 rd. ed. New York: John Wiley & Sons; 2003.
2. Selvin S. Statistical Analysis of epidemiologic data. 3 rd. ed. New York: Oxford University Press; 2004.
3. Pita Fernández S, Vila Alonso MT, Carpente Montero J. Determinación de factores de riesgo. Cad Aten Primaria 1997; 4: 75-78. [[Texto completo](#)]
4. Pita Fernández S, López de Ullibarri Galparsoro I. Número necesario de pacientes a tratar para reducir un evento. Cad Aten Primaria 1998; 96-98. [[Texto completo](#)]
5. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1991.
6. Armitage P, Berry G. Estadística para la investigación biomédica. Barcelona: Harcourt Brace; 1999.
7. Juez Martel P. Herramientas estadísticas para la investigación en Medicina y Economía de la Salud. Madrid: Centro de Estudios Ramón Areces; 2001.

Tabla 1. Tabla de contingencia general para la comparación de dos variables dicotómicas.

Característica B	Característica A		Total
	Presente	Ausente	
Presenta	a	b	a + b
Ausente	c	d	c + d
Total	a + c	b + d	n

Tabla 2. Tabla de contingencia para estudiar la asociación entre fumar durante la gestación y el bajo peso del niño al nacer. Estudio de seguimiento de 2000 gestantes.

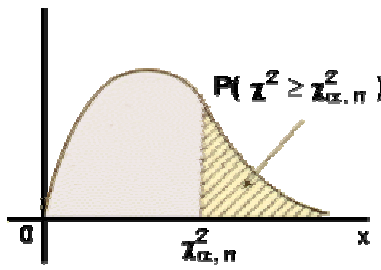
Gestante	Recién nacido de bajo peso		Total
	Sí	No	
Fumadora	43 (a)	207 (b)	250
No fumadora	105 (c)	1645 (d)	1750
Total	148	1852	2000

Tabla 3. Tabla de contingencia para estudiar la asociación entre fumar durante la gestación y el bajo peso del niño al nacer. Valores observados y valores esperados (entre paréntesis) si los factores fuesen independientes.

Gestante	Recién nacido de bajo peso		Total
	Sí	No	
Fumadora	43 (18.5)	207 (231.5)	250
No fumadora	105 (129.5)	1645 (1620.5)	1750

Total	148	1852	2000
-------	-----	------	------

Tabla 4. Distribución de ji-cuadrado



Grados de libertad	Probabilidad de un valor superior				
	0,1	0,05	0,025	0,01	0,005
1	2,71	3,84	5,02	6,63	7,88
2	4,61	5,99	7,38	9,21	10,60
3	6,25	7,81	9,35	11,34	12,84
4	7,78	9,49	11,14	13,28	14,86
5	9,24	11,07	12,83	15,09	16,75
6	10,64	12,59	14,45	16,81	18,55
7	12,02	14,07	16,01	18,48	20,28
8	13,36	15,51	17,53	20,09	21,95
9	14,68	16,92	19,02	21,67	23,59
10	15,99	18,31	20,48	23,21	25,19
11	17,28	19,68	21,92	24,73	26,76
12	18,55	21,03	23,34	26,22	28,30
13	19,81	22,36	24,74	27,69	29,82
14	21,06	23,68	26,12	29,14	31,32
15	22,31	25,00	27,49	30,58	32,80
16	23,54	26,30	28,85	32,00	34,27
17	24,77	27,59	30,19	33,41	35,72
18	25,99	28,87	31,53	34,81	37,16
19	27,20	30,14	32,85	36,19	38,58
20	28,41	31,41	34,17	37,57	40,00
21	29,62	32,67	35,48	38,93	41,40
22	30,81	33,92	36,78	40,29	42,80
23	32,01	35,17	38,08	41,64	44,18
24	33,20	36,42	39,36	42,98	45,56
25	34,38	37,65	40,65	44,31	46,93
26	35,56	38,89	41,92	45,64	48,29
27	36,74	40,11	43,19	46,96	49,65
28	37,92	41,34	44,46	48,28	50,99
29	39,09	42,56	45,72	49,59	52,34
30	40,26	43,77	46,98	50,89	53,67
40	51,81	55,76	59,34	63,69	66,77
50	63,17	67,50	71,42	76,15	79,49
60	74,40	79,08	83,30	88,38	91,95
70	85,53	90,53	95,02	100,43	104,21
80	96,58	101,88	106,63	112,33	116,32
90	107,57	113,15	118,14	124,12	128,30
100	118,50	124,34	129,56	135,81	140,17