

Comentarios de Bioestadística

El meta-análisis, parte I: fundamentos estadísticos

Gabriel Cavada Ch.^{1,2}¹Facultad de Medicina, Universidad de los Andes.²División de Bioestadística, Escuela de Salud Pública, Universidad de Chile.

The meta-analysis, Part I: statistical foundations

Un tema principal en la lectura y la investigación biomédica son los meta análisis, de hecho un meta análisis hecho sobre ensayos clínicos aleatorizados y enmascarados es el portador de información que entrega el mejor nivel de evidencia. Sin embargo, ¿qué es un meta análisis?, ¿Cuáles son los fundamentos teóricos que permiten la metodología?, son algunas de las preguntas que deberíamos contestar para entender este tipo de análisis.

La paradoja de Simpson

Un problema clásico que se presenta en el análisis estadístico puro es la llamada paradoja de Simpson, este fenómeno es cambio en el sentido de una asociación entre dos variables (numéricas o cualitativas) cuando se controla el efecto de una tercera variable:

Por ejemplo: En un estudio comparativo sobre tolerancia de dos fármacos antihipertensivos se determina la presencia o no de reacciones adversas al medicamento (RAM) leves, al codificar la información en Tratamiento (TTO = 0 ó TTO = 1) y RAM (ram = 0 sin RAM ó ram = 1 con RAM), la información se tabula a continuación:

ram	TTO		Total
	0	1	
0	434 82,67	410 78,10	844 80,38
1	91 17,33	115 21,90	206 19,62
Total	525 100,00	525 100,00	1,050 100,00

Como se observa el tratamiento 1 presenta una tasa de RAM de 21,9% y el tratamiento 0 una tasa de 17,3%, el test de independencia de ² sólo indica una asociación débil (no significativa) con un p-value = 0,062 y se estima un OR = 1,34.

Cuando esta información se desagrega por género, en hombres se obtiene:

ram	TTO		Total
	0	1	
0	351 86,67	122 93,85	473 88,41
1	54 13,33	8 6,15	62 11,59
Total	405 100,00	130 100,00	535 100,00

El tratamiento 1 presenta una tasa de RAM de 6,15% y el tratamiento 0 una tasa de 13,3%, el test de independencia de ² sólo indica una asociación significativa un p-value = 0,026 y se estima un OR = 0,43, es decir, el tratamiento 1 es protector al evento de observar una RAM.

Sin embargo para el caso de las mujeres se obtiene:

ram	TTO		Total
	0	1	
0	83 69,17	288 72,91	371 72,04
1	37 30,83	107 27,09	144 27,96
Total	120 100,00	395 100,00	515 100,00

El tratamiento 1 presenta una tasa de RAM del 27,09% y el tratamiento 0 una tasa de RAM del 30,83%, la asociación no fue significativa (p-value = 0,423) y un OR = 0,83.

Un análisis epidemiológico básico, nos indicaría la precaución de ajustar los resultados por sexo, pues hay una evidente diferencia. Esta diferencia nos lleva a dos conclusiones distintas: la primera es que en ambos sexos el tratamiento 1 protege de hacer RAM, sólo que en las mujeres falta tamaño de muestra para rechazar la hipótesis de independencia. O bien que el tratamiento 1 sólo protege de hacer una RAM a hombres. ¿Cómo saberlo? Si la segunda conclusión fuera la cierta habrían antecedentes clínicos, epidemiológicos o en las ciencias básicas que harían plausible la conclusión. Así es que supongamos que la primera conclusión es la correcta, en este escenario, habría que estimar una asociación (OR) entre

Comentarios de Bioestadística

RAM y tratamiento, ponderando dichas asociaciones por sus respectivos tamaños de muestras. Una forma de resolver este problema hacer esta estimación del OR mediante el método de Mantel y Haenszel, el que permite construir un OR más “potente” que los conseguidos en la estratificación, pero no se calculan a partir de la primera Tabla, en que no se distinguen los géneros. Los resultados de este OR combinado se muestran a continuación:

Sexo	OR	[95% Conf. Interval]	M-H Weight	
Masculino	,4262295	,1704908	,9360894	12,31402 (exact)
Femenino	,8334272	,5229791	1,344188	20,69126 (exact)
Crude	1,337711	,9737047	1,840389*	(exact)
M-H combined	,6815048	,4678835	,9926591**	

Test of homogeneity (M-H); $\chi^2(1) = 2,21$ Pr > $\chi^2 = 0,1375$.

Test that combined OR = 1: Mantel-Haenszel $\chi^2(1) = 3,85$
Pr > $\chi^2 = 0,0496$

El reglón* muestra el OR que se obtiene al combinar los datos mientras que el reglón** muestra el OR se que se obtiene al ponderar los OR por sus respectivos tamaños de muestras, cuando se opta por esta ponderación, la conclusión es que el tratamiento 1 protege de hacer una RAM tanto en hombres como en mujeres.

Este ejemplo, nos hace reflexionar en una metodología que combinara uno o más estudios, con resultados concluyentes (significativos) o no con el fin de obtener una conclusión más “potente”.

La distribución de probabilidades de un p-value

Un interesante resultado de la teoría de probabilidades para una variable continua, es que cualquiera sea el test estadístico que se utilice la transformación $-2\ln(F(p))$, donde p es el p-value esta tiene una distribución χ^2 con 2 grados de libertad. Para probar esta afirmación basta recordar que la distribución χ^2 con dos grados de libertad tiene como función densidad de probabilidades la expresión:

$$f(x) = \frac{1}{2} e^{-\frac{x}{2}}$$

Y que $\frac{dF}{dx} = f(x)$, con lo que es fácil seguir la siguiente cadena de implicaciones:

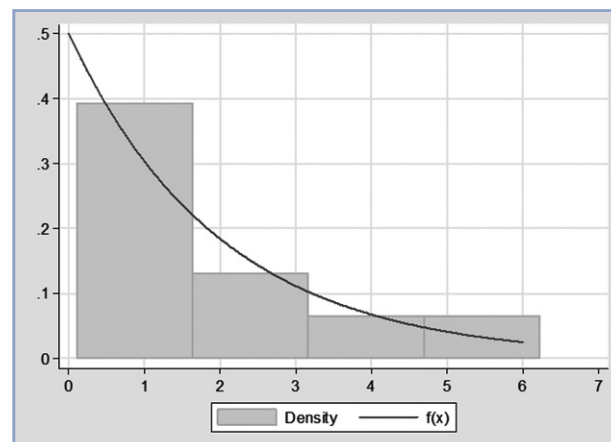
$$-2\ln(F(x)) = x \rightarrow \ln(F(x)) = -\frac{x}{2} \rightarrow F(x) = e^{-\frac{x}{2}} \rightarrow \frac{dF}{dx} = f(x) = \frac{1}{2} e^{-\frac{x}{2}}$$

Una verificación empírica de este resultado es la siguiente, suponga que se obtienen 20 p-values obtenidos de un

mismo test estadístico sobre muestras obtenidas de la misma población, como los tamaños de muestra de cada test son distintos, no necesariamente se obtienen p-values significativos, y estos podrían ser:

p-values
0,1370
0,6432
0,5578
0,6048
0,6842
0,1087
0,6185
0,0611
0,5552
0,8714
0,2551
0,0445
0,4242
0,8983
0,5219
0,8414
0,2110
0,5644
0,2648
0,9477

El siguiente gráfico muestra el histograma de la variable $-2\ln(p)$ y la curva corresponde a la función densidad de probabilidades de la χ^2 con dos grados de libertad:



Este hecho permite calcular la probabilidad de que un p-value sea mayor (o menor) que un determinado valor usando la transformación propuesta. Con ello podemos conocer la probabilidad de que un conjunto de estudios en que cada uno de ellos aporte un p-value, entregue un “meta-p-value” significativo y con ello una medida de asociación significativa.

En el próximo artículo se verá el cómo se hace operativo este resultado.