

NUEVAS INFERENCIAS ASINTÓTICAS SOBRE LA DIFERENCIA, RAZÓN Y COMBINACIÓN LINEAL DE DOS PROPORCIONES INDEPENDIENTES

Álvarez Hernández, M.¹ y Martín Andrés, A.²

^{1,2} Bioestadística. Facultad de Medicina. Universidad de Granada

RESUMEN

En medicina y estadística aplicada es muy habitual la necesidad de realizar inferencias asintóticas acerca de la diferencia ($d = p_2 - p_1$) o de la razón ($R = p_2/p_1$) de dos proporciones independientes. Ambos casos pueden enmarcarse dentro del caso general de efectuar inferencias acerca de una combinación lineal de dos proporciones $L = \beta_1 p_1 + \beta_2 p_2$ (un parámetro sobre el que recientemente se ha mostrado interés desde la perspectiva de los ensayos clínicos). En este trabajo se proponen y evalúan 10 nuevos métodos de inferencia para cada uno de los tres parámetros citados, concluyéndose que para errores bajos ($\alpha = 1\%$) alguno de los nuevos procedimientos tienen mejor comportamiento que los seleccionados por la literatura. En el caso L , el método óptimo es el basado en el test de razón de verosimilitudes (tras incrementar los datos en 0.5). En el caso d (caso R) el método óptimo consiste en sumar 0.5 (sumar 1) a todos los datos y aplicar una modificación de la transformación arco-seno.

Palabras y frases clave: Combinación lineal de dos proporciones; Diferencia de proporciones; Riesgo relativo; Test de la razón de verosimilitudes; Transformación arco-seno.

1. INTRODUCCIÓN

Las inferencias asintóticas sobre una combinación lineal $\Sigma \beta_i p_i$ de r proporciones binomiales independientes p_i han suscitado gran interés en los últimos años. El caso de $r \geq 3$ es el más novedoso (Newcombe, 2001; Price and Bonett, 2004; Tebbs and Roths, 2008; Zou et al., 2009, y Martín *et al.*, 2011a y b), pero el caso $r \leq 2$, ha estado presente en la literatura desde casi los inicios de la estadística. Cuando $r = 2$, el parámetro sobre el que hay que realizar la inferencia puede ser de tres tipos: la diferencia $d = p_2 - p_1$, la razón $R = p_2/p_1$ o una combinación lineal general $L = \beta_1 p_1 + \beta_2 p_2$.

Los casos d y R han sido profusamente tratados en la literatura -ver las referencias en Chan (2003), Dann and Koch (2005) y Martín and Álvarez (2012, 2013) entre otros- existiendo cientos de artículos en los que se analizan los diversos procedimientos. Por el contrario, el caso L ha recibido muy escasa atención -Phillips (2003) y Martín and Herranz (2010) para $\beta_1 < 0$ y Martín and Álvarez (2013) para cualquier valor de β_1 - aún cuando tiene la ventaja añadida de contener como caso particular a los casos d y R (por lo que cualquier propiedad del mismo se transmite a los otros dos). El trabajo actual está centrado en los parámetros L , d y R .

Las inferencias aludidas (que serán de dos colas) pueden referirse a la realización de un test de hipótesis sobre el parámetro implicado (hipótesis nulas $H_L: L = \lambda$, $H_d: d = \delta$ or $H_R: R = \rho$) o a la determinación de un intervalo de confianza para dicho parámetro (CI en adelante). Como el CI puede obtenerse por inversión del test y el test puede realizarse a través del CI (Agresti and Min, 2001), en la práctica es indiferente aludir a una u otra inferencia.

El objetivo de este trabajo es proponer nuevos métodos de inferencia asintóticos de dos colas para L , d y R , compararlos con los mejores métodos conocidos y seleccionar el mejor de todos ellos.

2. MÉTODOS PARA REALIZAR LA INFERENCIA

2.1. Métodos propuestos

Sean dos variables aleatorias binomiales independientes $x_i \sim B(n_i; p_i)$, en donde $i = 1$ y 2 , n_i es el tamaño de muestra y las proporciones p_i son desconocidas. Sea $L = \beta_1 p_1 + \beta_2 p_2$ el parámetro de interés sobre el que se desea realizar el test para la hipótesis nula $H_L: L = \lambda$. Cuando $\beta_1 = -1$ y $\beta_2 = +1$, se obtiene el test para $H_d: d = \delta$; cuando $\beta_1 = -\rho$, $\beta_2 = +1$ y $\lambda = 0$ se obtiene el test para $H_R: R = \rho$. Invirtiendo en λ , δ o ρ los tests anteriores se obtiene el CI para los parámetros L , d o R , respectivamente.

Sean $\bar{p}_i = x_i / n_i$, $\bar{L} = \sum \beta_i \bar{p}_i$, $\bar{d} = \bar{p}_2 - \bar{p}_1$ y $\bar{R} = \bar{p}_2 / \bar{p}_1$ las estimaciones muestrales de los parámetros poblacionales p_i , L , d y R , respectivamente. Hay muchos estadísticos disponibles para efectuar los tres tests anteriores, pero en este trabajo se propone realizar las inferencias en base a dos estadísticos cuya actuación no ha sido analizada hasta ahora. En primer lugar se propone el estadístico $4 \sum n_i (\sin^{-1} \sqrt{\bar{p}_i} - \sin^{-1} \sqrt{\hat{p}_i})^2$ que es una modificación del estadístico de Herranz and Martín (2008) basado en la transformación arco-seno. Para calcular el valor del mismo es preciso estimar las proporciones p_i . En el contexto de los parámetros d y R , Mee (1984) y Koopman (1984), respectivamente, proponen sustituir los valores desconocidos por sus estimadores de máxima verosimilitud $\hat{p}_i$ bajo H , obteniéndose así el siguiente estadístico AL:

$$z_{AL}^2 = 4 \sum n_i (\sin^{-1} \sqrt{\bar{p}_i} - \sin^{-1} \sqrt{\hat{p}_i})^2.$$

El segundo estadístico que se propone (notado por LR) es el del clásico test de razón de verosimilitudes (utilizado por Barker *et al.*, 2001, y Herranz and Martín, 2008, para el caso d):

$$z_{RL}^2 = -2 \left\{ x_1 \ln \frac{\hat{p}_1}{\bar{p}_1} + y_1 \ln \frac{1 - \hat{p}_1}{1 - \bar{p}_1} + x_2 \ln \frac{\hat{p}_2}{\bar{p}_2} + y_2 \ln \frac{1 - \hat{p}_2}{1 - \bar{p}_2} \right\}.$$

Una vez determinado el valor experimental del estadístico (z_{exp}^2 en general), el test se realiza del modo clásico: comparando z_{exp}^2 con $z_{\alpha/2}^2$, en donde $z_{\alpha/2}$ es el $(1-\alpha/2) \cdot 100$ percentil de una distribución normal estándar. Por inversión del test, resolviendo la ecuación $z_{exp}^2 = z_{\alpha/2}^2$ en el parámetro λ , δ or ρ , se obtiene el CI para L , d or R , respectivamente.

Es conocido que algunos métodos de inferencia asintóticos mejoran considerablemente por el simple hecho de incrementar los datos originales x_i y $y_i = n_i - x_i$ en una determinada cantidad h_i . Son tradicionales los incrementos $h_i = 0.5$ (Woolf, 1955, en el contexto del parámetro R) y $h_i = 1$ (Agresti and Caffo, 2000, en el contexto del parámetro d). Otras posibilidades de incrementos son los sugeridos por Martín *et al.* (2011a) en un contexto general.

Planteados los dos estadísticos (AL y LR), cada uno puede aplicarse con cada uno de los incrementos. El resultado es que se han definido 10 métodos de inferencia distintos para cada uno de los tres parámetros estudiados (L , d and R).

2.2. Métodos de inferencia seleccionados en la literatura

En el caso L , Martín and Álvarez (2013) centran su atención en la evaluación y comparación de distintos métodos de inferencia para $\beta_1 \neq -\beta_2$. Sus resultados indican que el método óptimo (que es de tipo adjusted Wald) consiste en utilizar el clásico estadístico W de Wald que sigue para los datos incrementados usualmente en una cantidad $z_{\alpha/2}^2 / 4$:

$$z_{W(L)}^2 = (\bar{L} - \lambda)^2 / \left(\frac{\beta_1^2 \bar{p}_1 \bar{q}_1}{n_1} + \frac{\beta_2^2 \bar{p}_2 \bar{q}_2}{n_2} \right).$$

En el caso d , Martín *et al.* (2012) evalúan comparativamente una gran cantidad de métodos, concluyendo que el óptimo es el método consistente en sumar 0.5 a todos los datos y utilizar el siguiente estadístico A basado en la transformación arco seno:

$$z_{A(d)}^2 = \frac{4n_1 n_2 (\bar{d}' - \delta')^2}{n_1 + n_2} \text{ donde } \bar{d}' = \sin^{-1} \sqrt{\bar{p}_2} - \sin^{-1} \sqrt{\bar{p}_1} \text{ y } \delta' = \sin^{-1} \sqrt{\hat{p}_2} - \sin^{-1} \sqrt{\hat{p}_1}.$$

En el caso R el método óptimo es el seleccionado por Martín and Álvarez (2012) consistente en sumar 0.5 a todos los datos y utilizar la siguiente aproximación S del estadístico de las marcas:

$$z_{S(R)}^2 = \begin{cases} \frac{n n_1 n_2 (\bar{p}_2 - \rho \bar{p}_1)^2}{(x_2 + \rho x_1) \{ (n_1 - x_2) + \rho (n_2 - x_1) \}} & \text{si } \frac{x_2}{n - x_1} \leq \rho \leq \frac{n - x_2}{x_1} \\ \frac{n_2 (\bar{p}_2 - \rho \bar{p}_1)^2}{\rho (1 - \rho)} & \text{si } \rho < \frac{x_2}{n - x_1} \\ \frac{n_1 (\bar{p}_2 - \rho \bar{p}_1)^2}{\rho - 1} & \text{si } \rho > \frac{n - x_2}{x_1}. \end{cases}$$

Sorprendentemente, otros métodos más complejos (incluido el muy reputado método de las marcas) presentan un peor comportamiento, especialmente en el caso de la diferencia y de la razón de proporciones (Chan, 1998; Newcombe, 1998; Santner et al., 2007; Martín et al., 2012; Martín y Álvarez, 2012).

3. EVALUACIÓN DE LOS MÉTODOS PROPUESTOS

Para cada uno de los 30 métodos a analizar hay que realizar los siguientes pasos:

- (1) Seleccionar uno de los errores $\alpha = 1\%, 5\%, 10\%$ y una de las parejas (n_1, n_2) con $n_1 \leq n_2$ y $n_i = 40, 60, 100$.
- (2) Seleccionar el valor de los parámetros λ, δ o ρ desconocidos (en todos los casos se utilizan los valores propuestos por los autores que se citan).
- (3) Construir la región crítica $CR = \{(x_1, x_2) \mid z_{exp}^2 \geq z_{\alpha/2}^2\}$ y el error real α^* (tamaño del test):

$$\alpha^* = \text{Máx}_{A < p_1 < B} \sum_{CR} \binom{n_1}{x_1} \binom{n_2}{x_2} p_1^{x_1} (1 - p_1)^{y_1} p_2^{x_2} (1 - p_2)^{y_2}, \text{ donde } A = \max \{0; (\lambda - \beta_2)/\beta_1\} \text{ y } B = \min \{1; \lambda/\beta_1\}$$

si $\beta_1 \beta_2 > 0$, $A = \max \{0; \lambda/\beta_1\}$ y $B = \min \{1; (\lambda - \beta_2)/\beta_1\}$ si $\beta_1 \beta_2 < 0$, y $p_2 = (\lambda - \beta_1 p_1) / \beta_2$.

Para evaluar la diferencia entre el error nominal (α) y el error real (α^*), se define el incremento $\Delta\alpha = 100(\alpha - \alpha^*)\%$.

- (4) Calcular el valor de $\theta = 100 \times (\text{número de puntos del conjunto } CR) / [(n_1 + 1)(n_2 + 1)]\%$, en donde $(n_1 + 1)(n_2 + 1)$ es el número total de puntos del espacio muestral.
- (5) Determinar si el método “falla”, es decir si $\Delta\alpha \leq -1\%, -2\%$ o -4% para $\alpha = 1\%, 5\%$ o 10% respectivamente, a fin de controlar el número de veces en las que el test es demasiado liberal.
- (6) Calcular el número de fallos y los valores medios de $\Delta\alpha$ y de θ en todas las combinaciones (n_1, n_2) para la gama de valores de cada parámetro de interés.

Una vez obtenidos los resultados, la selección del método óptimo se realizará bajo los siguientes criterios: (a) se dará especial importancia al caso $\alpha = 5\%$, (b) se descartarán los métodos con un excesivo número de fallos, (c) se eligen de ellos los que tengan un $\Delta\alpha$ más cercano a 0 y los de mayor potencia, (d) se preferirá el método que sea más sencillo de aplicar.

Finalmente se comparan los procedimientos preferidos con los seleccionados en la literatura, a fin de escoger el mejor método para cada uno de los parámetros.

REFERENCIAS

- Agresti, A. and Min, Y. (2001). On small-sample confidence intervals for parameters in discrete distributions. *Biometrics* 57, 963-971.
- Barker, L., Rolka, H., Rolka, D. and Brown, C. (2001). Equivalence testing for binomial random variables: which test to use?. *The American Statistician* 55, 279-287.
- Dann, R.S. and Koch, G.G. (2005). Review and evaluation of methods for computing confidence intervals for the ratio of two proportions and considerations for non-inferiority clinical trials. *Journal of Biopharmaceutical Statistics* 15, 85-107.

- Herranz Tejedor, I. and Martín Andrés, A. (2008). A numerical comparison of several unconditional exact tests in problems of equivalence based on the difference of proportions. *Journal of Statistical Computation and Simulation* 78 (11), 969-981.
- Koopman, P.A.R. (1984). Confidence intervals for the ratio of two binomial proportions. *Biometrics* 40, 513-517.
- Martín Andrés, A. and Álvarez Hernández, M. (2012). Two-tails approximate confidence intervals for the ratio of proportions. *Statistics and Computing*. Prepublished September, 2012.
- Martín Andrés, A. and Álvarez Hernández, M. (2013). Optimal method for realizing two-sided inferences about a linear combination of two proportions. *Communications in Statistics: Simulation and Computation* 42 (2), 327-343.
- Martín Andrés, A., Álvarez Hernández, M. and Herranz Tejedor, I. (2011). Inferences about a linear combination of proportions. *Statistical Methods in Medical Research* 20, 369-387.
- Martín Andrés, A., Álvarez Hernández, M. and Herranz Tejedor, I. (2012). Asymptotic two-tailed confidence intervals for the difference of proportions. *Journal of Applied Statistics* 39 (7), 1423-1435.
- Phillips, K.F. (2003). A new test of non-inferiority for anti-infective trials. *Statistics in Medicine* 22, 201-212.
- Price, R. M. and Bonett, D. G. (2004). An improved confidence interval for a linear function of binomial proportions. *Computational Statistics & Data Analysis* 45 (3), 449-456.
- Tebbs, J. M. and Roths, S. A. (2008). New large-sample confidence intervals for a linear combination of binomial proportions. *Journal of Statistical Planning and Inference* 138 (6), 1884-1893.
- Zou, G.; Huang, W. and Zhang, X. (2009). A note on confidence interval estimation for a linear function of binomial proportions. *Computational Statistics & Data Analysis* 53, 1080-1085.