INFERENCIAS ASINTÓTICAS SOBRE UNA COMBINACIÓN LINEAL DE *K* PROPORCIONES

María Álvarez Hernández*, Antonio Martín Andrés*, Inmaculada Herranz Tejedor** y Pedro Femia Marzo*.
*Bioestadística. Facultad de Medicina. Universidad de Granada (Spain)
**Bioestadística. Facultad de Medicina. Universidad Complutense de Madrid (Spain)

ABSTRACT

Statistical methods for carrying out asymptotic inferences (test or confidence intervals) relative to one or two independent binomial proportions are very frequent in the literature. However, inferences about a linear combination of more than two independent proportions $L = \Sigma\beta_i p_i$ have had very little attention paid to them (focused almost exclusively on the classic Wald method).
This paper reviews the study by Martín *et al.* [19, 20] for the performance of such inferences, mainly from the more efficient viewpoint of the score method. That article offers approximate formulas that are easy to calculate, gives a general proof of Agresti's heuristic method (consisting of adding a certain number of successes and failures to the original results before applying Wald's method), the previous more powerful methods are modified giving them a continuity correction and proves that the score method (which verifies the desirable properties of spatial convexity and parametric convexity) is the best option in comparison with other methods, which can be solved using a free programme which is available from the webpage http://www.ugr.es/local/bioest/Z_LINEAR_K.EXE.

KEY WORDS: Linear function of proportions; Score method; Peskun method; Wald method; Wilson method.

RESUMEN

Los métodos estadísticos para efectuar inferencias asintóticas (tests o intervalos de confianza) relativas a una o dos proporciones binomiales independientes son muy frecuentes en la literatura. Sin embargo las inferencias acerca de una combinación lineal de más de dos proporciones independientes $L = \Sigma\beta_i p_i$ han recibido escasa atención (centrada casi exclusivamente en el clásico método de Wald).
Esta monografía resume la investigación desarrollada por Martín *et al.* [19, 20] sobre este tipo de inferencias, principalmente desde el más eficiente punto de vista del método de las marcas. En dichos trabajos se ofrecen fórmulas aproximadas de cálculo sencillo, se proporciona una demostración general del método heurístico de Agresti (consistente en añadir a los datos originales un cierto número de éxitos y de fallos antes de aplicar el método de Wald), se dota a los métodos más potentes de una corrección por continuidad y se demuestra que el método de las marcas (que verifica las deseables propiedades de convexidad espacial y paramétrica) es óptimo frente a otros métodos, proporcionando un programa gratuito que permite su resolución (http://www.ugr.es/local/bioest/Z_LINEAR_K.EXE).

PALABRAS CLAVE: Funciones lineales de proporciones; método Score; método de Peskun ; método de Wald ; método de Wilson

1. INTRODUCCIÓN

En los últimos años han tomado gran importancia [37] las inferencias acerca de una combinación lineal $L=\Sigma\beta_i p_i$ de K proporciones binomiales independientes p_i (es decir, inferencias basadas en la toma de muestras independientes de las poblaciones objetivo), especialmente en el ámbito de las Ciencias de la Salud.

Muestras	SÍ	NO	Total	Coeficientes
1	x_1	y_1	n_1	β_1
2	x_2	y_2	n_2	β_2
…	…	…	…	…
K	x_K	y_K	n_K	β_K
Total	a_1	a_2	n	

Tabla 1

Tabla 2×*K* para muestras independientes

Los datos obtenidos en este tipo de estudios suelen presentarse en el formato de la Tabla 1, en donde SÍ/NO alude a la presencia o ausencia de la característica que se estudia, x_i (y_i) es el nº de individuos de entre los n_i (tamaño de muestra) que sí (no) presentan la característica, β_i son los coeficientes de la combinación lineal (que habitualmente son conocidos y distintos de 0), $a_1 = \Sigma x_i$ ($a_2 = \Sigma y_i$) son los totales de individuos que sí (no) presentan la característica y, finalmente, $n = a_1+a_2 = \Sigma n_i$ es el tamaño total de la experiencia. Las variables aleatorias (x_i) siguen distribuciones binomiales independientes $x_i \longrightarrow B(n_i, p_i)$, con $i = 1,2,...,K$, en donde p_i es la proporción (desconocida) de individuos de la población i que presentan la característica en estudio.

La combinación lineal L puede ser en unas ocasiones un contraste ($\Sigma\beta_i = 0$), en cuyo caso suele interesar realizar el test para H_0: $L = 0$ o determinar un intervalo de confianza para L (IC en adelante), el cual puede obtenerse por inversión del test para H_0: $L = 0$. En otras ocasiones la combinación lineal L puede no ser un contraste ($\Sigma\beta_i \neq 0$), en cuyo caso suele interesar determinar un IC para L, el cual puede obtenerse también mediante la inversión del test para H_0: $L = \lambda$.

Históricamente, la atención se ha centrado en los casos con $K \leq 2$, especialmente en el ámbito de la investigación médica. Ejemplos prácticos son el caso de los ensayos clínicos, que comparan dos tratamientos a través de la respuesta a una característica específica (usualmente el éxito del tratamiento) o el caso de los estudios epidemiológicos, que tratan de ver la relación entre un factor de riesgo y una enfermedad o efecto indeseado. Cuando $K = 2$, los objetivos pueden ser varios:

- Realizar inferencias sobre la diferencia $d = p_2–p_1$ de dos proporciones (es decir, sobre L para $\beta_1 = -1$ y $\beta_2 = +1$): Newcombe [27], Agresti and Caffo [2], Kang and Chen [16], Martín and Herranz [21], Brown and Li [6], Santner et al. [34], etc.
- Realizar inferencias sobre la razón $R = p_2/p_1$ de dos proporciones (es decir, sobre $L = p_2–Rp_1$ para $\beta_1 = –R$ y $\beta_2 = +1$): Chan [8], Agresti [1], Dann and Koch [12], Price and Bonnet [31], etc.
- Realizar inferencias sobre una combinación lineal de dos proporciones $L = \beta_2 p_2+\beta_1 p_1$ (con $\beta_1 < 0$): Phillips [30] y Martín and Herranz [24].

Cuando $K = 1$ (con lo que puede hacerse $\beta_1 = 1$), el objetivo será realizar inferencias sobre la única proporción p_1: Agresti and Coull [3], Newcombe [27], Brown *et al* [7], Agresti and Caffo [2].

Los casos con $K > 2$ son bastante menos habituales pero, aunque en los últimos años se les está prestando más atención dado su gran interés práctico [28, 31, 35, 37, 4, 39], hasta muy recientemente el problema prácticamente solo se ha abordado desde el punto de vista de los IC obtenidos por el clásico método de Wald (que Price and Bonett [34] y Schaarschmidt *et al.* [35] mejoran mediante el incremento de los datos en un cierto número de éxitos y de fracasos, lo que da lugar a los métodos adjusted Wald). Otro opción planteada por la literatura es la del método de Newcombe [28] que este utilizó para $K = 4$ y $\sum\beta_i = 0$ y que Zou *et al.* [39] generalizó para cualquier valor de K y de $\sum\beta_i$: sustituir las proporciones del método de Wald por los extremos del IC de Wilson [38] para cada una de ellas.

El objetivo de este trabajo es hacer una revisión del análisis desarrollado por Martín *et al.* [19, 20] para las inferencias asintóticas acerca de una combinación lineal de varias proporciones. En dicha investigación se evalúan los métodos planteados en la literatura durante los últimos años y se proponen nuevos procedimientos como el método de las marcas (sobre el que hay común acuerdo de que produce mejores resultados que el de Wald, incluso para cualquier parámetro de una tabla de contingencia [18]) y el método de Peskun (basado en el criterio de Sterne [36]). Adicionalmente, el análisis del método de las marcas proporciona una prueba teórica del resultado heurístico de que el intervalo de confianza de Wald al 95% mejora si a cada muestra se le añaden $2/K$ éxitos y fallos, al tiempo que se generaliza el resultado para cualquier valor de la confianza. Finalmente se plantea una posible mejora de cualquier método si al mismo se les dota de una corrección por continuidad (cc en adelante) y se comparan los métodos propuestos (mediante estudios de simulación) con el fin de seleccionar el mejor de ellos.

2. GENERALIDADES SOBRE EL TEST

Sean K variables aleatorias binomiales independientes y $L = \Sigma\beta_i p_i$ el parámetro de interés (con las proporciones p_i desconocidas y los parámetros β_i conocidos). Como el estadístico $\bar{L} = \Sigma\beta_i \bar{p}_i$, con $\bar{p}_i = x_i/n_i$, es asintóticamente normal con media $L = \Sigma\beta_i p_i$ y varianza $\Sigma\, \beta_i^2\, p_i q_i/n_i$, siendo $q_i = 1–p_i$, entonces para contrastar H_0: $L = \lambda$ *vs.* H_1: $L \neq \lambda$ (en donde $\sum_{\beta_i<0}\beta_i = B^- \leq \lambda \leq B^+ = \sum_{\beta_i>0}\beta_i$, con $B = \Sigma\beta_i$ y $n = \Sigma n_i$) hay que comparar del modo clásico el estadístico

$$z_{exp}^2 = \frac{\left(\bar{L} - \lambda\right)^2}{\sum \beta_i^2 p_i q_i / n_i} \tag{1}$$

con $z_{\alpha/2}^2$ (en donde $z_{\alpha/2}$ es el percentil $1–\alpha/2$ de la distribución normal típica). La inversión del test, haciendo $z_{exp}^2 = z_{\alpha/2}^2$ y despejando λ, proporciona el IC al $(1–\alpha)\times100\%$ de confianza para L. Estas expresiones no tienen utilidad hasta que las proporciones p_i desconocidas sean sustituidas por alguna estimación de las mismas.

2.1. Método de Wald y método de Newcombe

La opción más simple y conocida consiste en sustituir p_i por $\bar{p}_i$ en la expresión (1). Esto da lugar a los clásicos estadísticos de Wald y IC de Wald siguientes (en donde $\bar{q}_i = 1 - \bar{p}_i$):

$$z_{\mathrm{W}}^2 = \frac{\left(\bar{L} - \lambda\right)^2}{\sum \beta_i^2 \bar{p}_i \bar{q}_i / n_i},\ \mathrm{IC_W}\!: \bar{L} \pm z_{\alpha/2}\sqrt{\sum \beta_i^2 \bar{p}_i \bar{q}_i / n_i}, \tag{2}$$

obteniendo así lo que en adelante se denominará como procedimiento W.

Otra opción algo más complicada es la de Zou *et al.* [39], quienes justifican teóricamente y generalizan el procedimiento propuesto por Newcombe [27, 28] para ciertos casos especiales de $K = 2$ y $K = 4$. El procedimiento (denominado en adelante por procedimiento N) consiste en sustituir las proporciones desconocidas p_i por el extremo apropiado $\ddot{p}$ del IC de Wilson [38] para las mismas:

$$\ddot{p} = \begin{cases} u_i\ (si\ \beta_i > 0),\ \ l_i\ (si\ \beta_i < 0) cuando\ \bar{L} < \lambda \\ l_i\ (si\ \beta_i > 0),\ \ u_i\ (si\ \beta_i < 0) cuando\ \bar{L} > \lambda \end{cases}$$

con

$$l_i = \frac{x_i + \frac{z_{\alpha/2}^2}{2} - z_{\alpha/2}\sqrt{\frac{z_{\alpha/2}^2}{4} + \frac{x_i y_i}{n_i}}}{n_i + z_{\alpha/2}^2} \ \text{ y } \ u_i = \frac{x_i + \frac{z_{\alpha/2}^2}{2} + z_{\alpha/2}\sqrt{\frac{z_{\alpha/2}^2}{4} + \frac{x_i y_i}{n_i}}}{n_i + z_{\alpha/2}^2},$$

con lo cual (el IC que sigue es la expresión utilizada por Martín *et al*. y no la original de Zou *et al.*):

$$z_{\mathrm{N}}^2 = \begin{cases} \left(\bar{L} - \lambda\right)^2 / R^2(+) & \text{si } \bar{L}<\lambda \\ \left(\bar{L} - \lambda\right)^2 / R^2(-) & \text{si } \bar{L}>\lambda \end{cases},\ IC_{\mathrm{N}}\!: \begin{cases} L_1 = \bar{L} - z_{\alpha/2} R(-) \\ L_2 = \bar{L} + z_{\alpha/2} R(+) \end{cases}, \tag{3}$$

en donde:

$$R(+)=\sqrt{\sum_{\beta_i>0} \frac{\beta_i^2 u_i (1-u_i)}{n_i} + \sum_{\beta_i<0} \frac{\beta_i^2 l_i (1-l_i)}{n_i}},\ R(-)=\sqrt{\sum_{\beta_i>0} \frac{\beta_i^2 l_i (1-l_i)}{n_i} + \sum_{\beta_i<0} \frac{\beta_i^2 u_i (1-u_i)}{n_i}}.$$

Los dos procedimientos anteriores están basados en estimadores de las proporciones p_i que no están restringidos por la hipótesis nula H_0: $L = \lambda$.

2.2. Método de las marcas

Martín *et al.* [19] proponen sustituir las p_i por sus estimadores de máxima verosimilitud $\hat{p}_i$ bajo H_0, obteniendo el IC por inversión del test. Esto da lugar al procedimiento S (que, como demostraron dichos autores, es equivalente al método de las marcas) consistente en resolver en C, λ o z_{S}^2 (según sea el objetivo) la ecuación:

$$y = n + (B–2\lambda)C – \Sigma R_i = 0 \tag{4}$$

en donde $C = z_{\mathrm{S}}^2/(\bar{L} – \lambda)$, $R_i^2 = n_i^2 + \beta_i^2 C^2 + 2 n_i b_i \beta_i C$ y $b_i = 1–2\,\bar{p}_i$.

Cuando el objetivo es realizar el test (en cuyo caso λ es conocido) y $\bar{L} \neq \lambda$, entonces el estadístico z_{S}^2 de las marcas es la única solución $z_{\mathrm{S}}^2 \neq 0$ de la ecuación (4); cuando $\bar{L} = \lambda$ se asume que $z_{\mathrm{S}}^2 = 0$. Alternativamente, si el investigador no desea conocer el valor de z_{S}^2, sino sólo saber si el test es significativo al error α, entonces basta aplicar la siguiente regla:

$$\text{Decidir } H_1 \Leftrightarrow y\left(C = z_{\alpha/2}^2 / \left(\bar{L} - \lambda\right)\right) \geq 0 \tag{5}$$

lo que simplifica los cálculos enormemente (su intensidad de cálculos es similar a la del test de Wald). Adicionalmente, si el objetivo es obtener los estimadores de máxima verosimilitud bajo H_0 entonces la ecuación se resuelve en C y $\hat{p}_i = (n_i+\beta_i C-R_i)/2\beta_i C$.

Cuando el objetivo es obtener el IC $L_I \le L \le L_S$ (en cuyo caso $z_S^2 = z_{\alpha/2}^2$ es conocido), entonces L_i son las únicas dos soluciones λ de la ecuación (4). Si no existe solución para el extremo inferior (superior), entonces $L_I = B^-$ ($L_S = B^+$). Similarmente cuando el objetivo sea obtener el IC para β_K en valores fijados de λ, $\beta_i \ne \beta_K$ y $z_S^2 = z_{\alpha/2}^2$.

Obsérvese que todo lo anterior contiene como resultados particulares a los casos de una proporción ($K = 1$), de la diferencia de dos proporciones ($L = d = p_2 - p_1$) y de la razón de riesgos ($L = p_2 - Rp_1$ y $\lambda = 0$). En particular los tests e IC de Mee [25] para d y de Koopman [17] para R son casos particulares del caso general L. De igual modo, la demostración de que $z_S^2 = \chi^2$ que proporcionan los autores, contiene como casos particulares a las demostraciones de Nam [26] y Gart and Nam [13] para d y R respectivamente. De hecho, la expresión (4) fue demostrada por Martín and Herranz [23] para el caso d.

2.2.1. Aproximaciones generales y de tipo adjusted Wald.

Con el fin de simplificar la resolución de la ecuación (4) en z_S^2 (si se trata del test) o en λ (si se trata del IC), conviene obtener expresiones aproximadas de tal ecuación. Como demostraron Martín *et al.* [19], desarrollando en serie de Maclaurin el término R_i, la expresión (4) se convierte en esta otra:

$$(\bar{L}-\lambda)^3 - z_S^2(\bar{L}-\lambda)V_1 + z_S^4 V_2 \approx 0\text{, donde } V_1 = \sum\frac{\beta_i^2\bar{p}_i\bar{q}_i}{n_i},\ V_2 = \sum\frac{\beta_i^3\bar{p}_i\bar{q}_i b_i}{n_i^2}. \tag{6}$$

Si se retienen sólo los términos de $O(n_i) \ge -1$ y se divide por $(\bar{L}-\lambda)$ se obtienen las clásicas soluciones de Wald de la expresión (2). Si se retienen sólo los términos de $O(n_i) \ge -2$, se sustituye $z_S^4 \approx z_S^2(\bar{L}-\lambda)^2/V_1$ y se divide por $(\bar{L}-\lambda)$, entonces se obtiene $0 \approx (\bar{L}-\lambda)^2 V_1 + (\bar{L}-\lambda)V_2 z_S^2 - V_1 z_S^2$. De esto se deducen los siguientes estadístico e IC aproximados:

$$z_{S'}^2 = \frac{(\bar{L}-\lambda)^2}{V_1-(\bar{L}-\lambda)V_2/V_1},\ IC_{S'}: \bar{L} + z_{\alpha/2}^2\frac{V_2}{2V_1} \pm z_{\alpha/2}\sqrt{V_1 + z_{\alpha/2}^2\left(\frac{V_2}{2V_1}\right)^2} \tag{7}$$

Es conocido que los métodos heurísticos adjusted Wald tienen su origen en la propuesta de Agresti and Coull [3] para el caso de una proporción (K=1) en el que demostraron que el centro del IC de Wilson (que es el IC de las marcas para una proporción) es igual al centro del IC adjusted Wald cuando los datos son incrementados en $z_{\alpha/2}^2/2$, siendo esta la razón de la buena actuación de este último. En base a la aproximación (7), Martín *et al.* [19] demostraron que eso mismo es lo que aproximadamente sucede en el caso de $K > 1$, ya que el método adjusted Wald con un incremento de los datos de:

$$c_i = \frac{n_i z_{\alpha/2}^2}{2(Kn_i - z_{\alpha/2}^2)} = \frac{n_i h}{n_i - 2h} \approx h \text{ en donde } h = \frac{z_{\alpha/2}^2}{2K} \tag{8}$$

proporciona un IC cuyo centro es aproximadamente igual que el del $IC_{S'}$ de la expresión (7). Obsérvese que haciendo $c_i \approx h$ se obtiene el método adjusted Wald con un incremento de $2/K$ (que es el que proponen Price and Bonett [31] para α=5% si se considera que $z_{\alpha/2}^2 = 1{,}96^2 \approx 4$).

Las aproximaciones anteriores son correctas [19] solo cuando los datos observados no pertenecen a la frontera del espacio muestral, es decir cuando $0 < x_i < n_i$ ($\forall i$). En otro caso, estos autores demuestran que las soluciones correctas son mucho más complejas, pudiéndoselas simplificar en un método adjusted Wald basado en el incremento:

$$h_i' = \frac{z_{\alpha/2}^2}{2}\cdot\frac{1+\delta_{ij}K}{K} \tag{9}$$

en donde $\delta_{i1} = 1$ en todos los subíndices i que verifican { $\bar{p}_i = 0$ y $\beta_i < 0$} o { $\bar{p}_i = 1$ y $\beta_i > 0$} ($\delta_{i1} = 0$ en otro caso) y $\delta_{i2} = 1$ en todos los subíndices i que verifican { $\bar{p}_i = 0$ y $\beta_i > 0$} o { $\bar{p}_i = 1$ y $\beta_i < 0$} ($\delta_{i2} = 0$ en otro caso). Obsérvese que cuando $0 < \bar{p}_i < 1$ ($\forall i$) ambos incrementos (h y h_i') coinciden.

2.2.2. Propiedades de convexidad

Para que un estadístico T cualquiera sea útil en la inferencia es preciso que verifique ciertas propiedades de coherencia.

Barnard [5] indicó la conveniencia de que sean convexas las regiones críticas para el test clásico H_0: d = 0, lo que implica que el estadístico T debe ser creciente (decreciente) en $\bar{p}_2$ ($\bar{p}_1$), aunque tal crecimiento o decrecimiento no tiene que ser estricto. Röhmel and Mansmann [32] justificaron que lo mismo debe suceder en el caso más general de H_0: $d = \delta$. En el caso actual (H_0: $L = \lambda$) debería ocurrir que T sea creciente (decreciente) en las $\bar{p}_i$ con $\beta_i > 0$ ($\beta_i < 0$) –es decir que T verifique la propiedad de *convexidad espacial*– lo que implica que las regiones críticas no presentarán huecos. Martín *et al.* [19] demostraron que el estadístico z_S^2 verifica tal propiedad, demostración que contiene como casos particulares al caso de la diferencia H_0: $d = p_2 - p_1 = \delta$ (demostrado por Martín and Herranz [22]) y al caso de la razón de riesgos H_0: $p_2 - Rp_1 = p_2 - \rho p_1 = 0$.

Röhmel and Mansmann [33] señalaron la conveniencia de que el p-value $P(\delta)$ para el test H_0: $d = p_2 - p_1 \leq \delta$ sea creciente en δ. De modo general, para que el test H_0: $L = \lambda$ basado en T sea coherente es preciso que su p-value $P(\lambda)$ sea creciente (decreciente) en λ cuando $\lambda < \bar{L}$ ($\lambda > \bar{L}$). Esto implica que el estadístico debe ser decreciente en λ, es decir T debe verificar la propiedad de *convexidad paramétrica* en λ. La verificación de tal propiedad, que es la que garantiza que la inversión del test mediante la igualdad $T^2 = z_{\alpha/2}^2$ es equivalente a resolver la desigualdad $T^2 \leq z_{\alpha/2}^2$, proporciona un IC para λ que no presenta huecos. De modo similar, para que el IC para β_i sea coherente es preciso que T sea creciente en β_i (convexidad paramétrica en β_i). En el Anexo de Martín *et al.* [19], se demuestra igualmente que z_S^2 verifica las dos propiedades de convexidad paramétrica (lo que contiene como caso particular al caso d de Martín and Herranz [22]).

La anterior puede resumirse indicando que cualquier estadístico T debe de verificar las siguientes propiedades (z_S^2 las verifica):

$$\frac{dT}{d\bar{p}_i}\begin{cases}\geq 0 \ si\ \beta_i > 0\\ \leq 0\ si\ \beta_i < 0\end{cases},\quad \frac{dT}{d\lambda} \leq 0\ ,\quad \frac{dT}{d\beta_i} \geq 0 \tag{10}$$

2.3 Método de Peskun

El argumento que sigue (basado en el criterio de Sterne [36] y utilizado por Peskun [29] para el caso de la diferencia de proporciones) fue empleado por Martín *et al.* [20], con el fin de proponer el procedimiento P basado en una nueva estimación $\breve{p}_i$ de las proporciones p_i bajo H_0. El test de las marcas para H_0: $L = \lambda$ (con $L = \Sigma\beta_i p_i$) será significativo si $z_S^2 \geq z_{\alpha/2}^2$ en todos los valores p_i tales que $L = \lambda$; en consecuencia el objetivo debe ser determinar el mínimo valor de z_S^2 bajo la condición $\Sigma\beta_i p_i = \lambda$. Dichos autores demostraron que el estadístico y el IC son, respectivamente:

$$z_P^2 = \frac{4\left(\bar{L}-\lambda\right)^2}{\sum\frac{\beta_i^2}{n_i} - \frac{\left(B-2\lambda\right)^2}{n}}\ ,\ \mathrm{CI_P}{:}\ \frac{n}{n+z_{\alpha/2}^2}\left\{\bar{L} + \frac{Bz_{\alpha/2}^2}{2n} \pm \frac{z_{\alpha/2}}{2}\sqrt{\frac{n+z_{\alpha/2}^2}{n}\left(\sum\frac{\beta_i^2}{n_i}\right) - \frac{\left(B-2\bar{L}\right)^2}{n}}\right\} \tag{11}$$

3. INCREMENTO EN LOS DATOS ORIGINALES

Como ya se comentó en la sección 2.2.1, Price and Bonett [31], motivados por los resultados de Agresti and Coull [3] y Agresti and Caffo [2] para los casos de una proporción y de la diferencia de proporciones respectivamente, comprueban que el IC de Wald mejora sustancialmente si la expresión (2) se obtiene en base a los datos x_i+h_i, y_i+h_i y n_i+2h_i, con $h_i = 2/K$, es decir si a los datos originales se les añaden $2/K$ éxitos y $2/K$ fracasos. Esto da lugar al método adjusted Wald que denotaremos como W1 (en contraposición al método de Wald original que denotaremos como W0).

Más recientemente, Martín *et al.* [19] comprueban que la mejor opción del tipo adjusted Wald es la proporcionada por dicho método con el incremento siguiente:

$$h_i = \begin{cases} \dfrac{z^2_{\alpha/2}(1+I_iK)}{2K} \text{ si } \bar{L} > \lambda \\ \dfrac{z^2_{\alpha/2}(1+S_iK)}{2K} \text{ si } \bar{L} < \lambda \end{cases} \text{en donde} \begin{cases} I_i = \begin{cases} 1 \text{ si } \bar{p}_i = 0 \text{ y } \beta_i < 0 \text{ o } \bar{p}_i = 1 \text{ y } \beta_i > 0 \\ 0 \text{ en otro caso} \end{cases} \\ S_i = \begin{cases} 1 \text{ si } \bar{p}_i = 0 \text{ y } \beta_i > 0 \text{ o } \bar{p}_i = 1 \text{ y } \beta_i < 0 \\ 0 \text{ en otro caso} \end{cases} \end{cases} \quad (12)$$

(al que denotaremos como W3) aunque una buena alternativa es el método W2 basado en el incremento $h_i = z^2_{\alpha/2}/2K$. Obsérvese que cuando $0 < x_i < n_i$ ($\forall i$) entonces W2 = W3, y que si además se elige α = 5% entonces $z^2_{\alpha/2}/2K \approx 2/K$ y W1 > W2 = W3. Cuando $x_i = 0$ o $x_i = n_i$ en algún valor de i (es decir, cuando algún dato observado se encuentra en la frontera del espacio muestral), los cálculos del método W3 se complican pues entonces el valor de h_i es diferente según que se vaya a determinar el extremo inferior L_I (caso de $\bar{L} > \lambda$) o el extremo superior L_S (caso de $\bar{L} < \lambda$).

Se ve pues que el procedimiento W proporciona cuatro métodos (*W0, W1, W2* and *W3)* basados en los incrementos $h_i = 0$, $2/K$, $z^2_{\alpha/2}/2K$ y el indicado en la expresión (11), respectivamente. Lo mismo puede hacerse con los otros tres procedimientos N, S y P, obteniéndose así un total de 16 métodos de inferencia posibles: *W0,..., W3, N0,..., N3, S0,..., S3, P0,…, P2 y P3.*

4. MÉTODOS CON CORRECCIÓN POR CONTINUIDAD

Es conocida la conveniencia [10] de efectuar una cc cuando la distribución de una variable aleatoria discreta (como es la variable x_i) se aproxima a través de una variable continua (como es la variable normal). Haber [14] propone que una cc debe consistir en sumar o restar a la variable la mitad de su salto promedio. En este caso, la variable es $\bar{L}$ (el estadístico de contraste) y, como $\sum_{\beta_i<0}\beta_i \leq \bar{L} \leq \sum_{\beta_i>0}\beta_i$ (pues $0 \leq \bar{p}_i \leq 1$), su salto total será $\Sigma|\beta_i|$ y la mitad de su salto promedio, es decir la cc, será $c = \left(\sum|\beta_i|\right)/2(N-1)$, con $N = \prod(n_i+1)$, pues N es el número total de puntos $(x_1,.., x_K)$ del espacio muestral [20].

Para determinar el estadístico z^2_{exp} de la expresión (1) con cc basta con redefinirlo así: $z^2_{exp_c} = 0$ si $|\bar{L}-\lambda| \leq c$ y $z^2_{exp_c} = \left(|\bar{L}-\lambda| - c\right)^2 / \sum\beta_i p_i q_i / n_i$ si $|\bar{L}-\lambda| > c$. Esto ocasiona los nuevos estadísticos $z^2_{W_c}$, $z^2_{N_c}$ y $z^2_{P_c}$ obtenidos a través de las expresiones (2), (3) y (10) respectivamente. En el caso del test de las marcas, Martín *et al.* (2012) justifican que el estadístico $z^2_{S_c}$ se obtiene cambiando el valor z^2_S de la expresión (4) por el valor $z^2_{S_c}\left\{(\bar{L}-\lambda)/\left(|\bar{L}-\lambda|-c\right)\right\}^2$, siendo $z^2_{S_c}$ la incógnita de tal ecuación.

Similarmente, para determinar el *IC* con cc basta con invertir los tests anteriores, lo que ocasiona los nuevos intervalos IC_{W_c}, IC_{N_c} y IC_{P_c} obtenidos a través de las expresiones (2), (3) y (10) respectivamente. En el caso del IC_{S_c} de las marcas basta cambiar el valor z^2_S de la expresión (4) por $z^2_{\alpha/2}\left\{(\bar{L}-\lambda)/\left(|\bar{L}-\lambda|-c\right)\right\}^2$ y determinar sus dos soluciones λ_i con $\sum_{\beta_i<0}\beta_i \leq \lambda_I < \bar{L}-c$ y $\bar{L}+c < \lambda_S \leq \sum_{\beta_i>0}\beta_i$.

Sea cual sea el caso (IC o test), la introducción de la cc da lugar a cuatro nuevos procedimientos Wc, Nc, Sc y Pc provenientes de los métodos W, N, S y P respectivamente. Previsiblemente, por la propia definición de c, su efecto será despreciable para $K > 2$.

5. ESTUDIO Y SELECCIÓN DEL MÉTODO ÓPTIMO

Los métodos de inferencia que Martín *el al.* [19, 20] evalúan son inicialmente los 16 indicados (los métodos W0, W1, W2, W3, N0,..., P3), lo que incluye las propuestas más relevantes de la literatura; de ellos, 13 son métodos nuevos (los denominados por W2, W3, N1, N2, N3, E0, E1, E2, E3, P0, P1, P2, P3). En todo caso, el objetivo perseguido por los autores era seleccionar el método o métodos óptimo/s (bajo los criterios que se especificarán) y, una vez realizada esa selección inicial, evaluar comparativamente los métodos más potentes con los métodos con cc a que dan lugar.

5.1. Descripción del estudio de simulación y criterios para la selección del método óptimo.

Para efectuar la evaluación anterior se realiza un estudio de simulación [20] en cada una de las siguientes combinaciones de los parámetros (α, K, n_i, β_i):

- α = 5% (aunque, en ocasiones especiales, también se contemplarán los errores del 1% y del 10%).
- K = 3, (n_1, n_2, n_3) = (10, 10, 10), (30, 30, 30), (30, 10, 10) y (30, 20, 10), (β_1, β_2, β_3) = (1/3, 1/3, 1/3), (1, −1/2, −1/2), (−1, 1/2, 2) y (1, 1, −1).
- K = 4, (n_1, n_2, n_3, n_4) = (10, 10, 10, 10), (20, 20, 20, 20), (20, 20, 10, 10) y (20, 15, 10, 5), (β_1, β_2, β_3, β_4) = (1/4, 1/4, 1/4, 1/4), (−1, +1, −1, +1), (1/3, 1/3, 1/3, 1) y (−3, −1, 1, 3).

El proceso de simulación consistirá en lo siguiente:

1) Fijar una combinación (α, K, n_i, β_i, método a evaluar).

2) Para cada punto del espacio muestral (x_1, x_2,…, x_K), obtener el IC (λ_I; λ_S) para L al (1–α)·100% de confianza.

3) Generar K valores de una distribución uniforme [0, 1] -los cuales formarán el vector de proporciones (p_1, p_2,…, p_K)- y calcular el valor real de L para el mismo ($L = \Sigma\beta_i p_i$).

4) Calcular el recubrimiento R y la longitud l del IC del método a partir de las expresiones siguientes:

$$\begin{aligned} R &= \sum_{x_1=0}^{n_1}\sum_{x_2=0}^{n_2}\ldots\sum_{x_K=0}^{n_K}\prod_{i=1}^{K}\binom{n_i}{x_i}p_i^{x_i}q_i^{n_i-x_i}\,I\ \ x_1,x_2,\ldots,x_K \\ l &= \sum_{x_1=0}^{n_1}\sum_{x_2=0}^{n_2}\ldots\sum_{x_K=0}^{n_K}\prod_{i=1}^{K}\binom{n_i}{x_i}p_i^{x_i}q_i^{n_i-x_i}\ \ L_S-L_I \end{aligned} \tag{13}$$

donde $I(x_1, x_2, ..., x_K) = 1$ si el IC (L_I, L_S) que ocasionan las observaciones (x_1, x_2, ..., x_K) contiene a $L = \Sigma\beta_i p_i$ e $I(x_1, x_2, ..., x_K) = 0$ en otro caso.

5) El proceso se repite 10.000 veces, obteniendo así 10.000 valores R_j y l_j a partir de los cuales se determinan el recubrimiento medio (*Rmean*), el recubrimiento mínimo (*Rmin*), la longitud media (*lmean*) y el porcentaje de "fallos" del método, es decir, el porcentaje de en los que el recubrimiento es menor del 93% (R < 93).

Para seleccionar el método óptimo, se asumen las siguientes reglas de actuación:

(a) El primer paso será eliminar aquellos métodos que sean demasiado liberales, es decir los que tengan un valor excesivo del parámetro *R < 93* (pues no tiene interés un método que tenga muchos "fallos").

(b) De los métodos que resten, se seleccionará aquellos que, teniendo un valor de *R<93* pequeño, su valor de *Rmean* sea próximo al 95%, prefiriendo los métodos conservadores (*Rmean* > 95%) sobre los liberales (*Rmean* < 95%).

(c) De entre los métodos que resten, se seleccionarán aquellos en los que *Rmin* sea grande (y cercano al nominal) y *lmean* sea pequeño.

(d) Finalmente, si sigue habiendo más de un método seleccionado, se preferirá aquel que sea más sencillo de aplicar (es decir, el que requiera de menos cálculos).

5.2. Selección del método óptimo

A partir de los resultados obtenidos por los estudios de simulación (planteados como se indica en la sección anterior), Martín *et al.* [20] deducen que:

i) Si los tamaños de muestra son pequeños, es decir $n_i \leq 10$ ($\forall i$), el mejor método es el método W3.

ii) En otro caso, S0 es el mejor, pero una alternativa mucho más sencilla es el método W3 (aunque es algo conservador, tiene algún fallo y provoca unos IC algo más amplios que S0);

iii) Si se desea un método que no falle nunca, el método P0 es el mejor (pero es demasiado conservador y proporciona IC excesivamente amplios);

iv) Si se desea utilizar siempre un mismo método, la mejor opción es el método S0c.

Por tanto, las fórmulas aconsejadas para realizar inferencias serán las siguientes:

- *Método S0 (el mejor método, salvo que todos los tamaños muestrales sean menores o iguales que 10)*

La expresión base es:

$$y = n\ \ \bar{L}-\lambda\ \ +\ \ B-2\lambda\ \ z-\left[\text{Signo}\ \ \bar{L}-\lambda\ \ \right]\sum R_i = 0$$

con $R_i^2 = n_i^2 \left(\bar{L}-\lambda\right)^2 + \beta_i^2 z^4 + 2n_i\beta_i\left(1-2\bar{p}_i\right) z^2\left(\bar{L}-\lambda\right)$ y $B^- = \sum_{\beta_i<0}\beta_i \le \lambda_I < \bar{L} < \lambda_S \le B^+ = \sum_{\beta_i>0}\beta_i$.

Para obtener el intervalo (en cuyo caso λ es desconocido), basta hacer en la misma $z = z_{\alpha/2}$ y obtener las dos soluciones (λ_I ; λ_S) de la ecuación en λ. Para obtener el estadístico de contraste para el test (en cuyo caso λ es conocido), basta obtener la única solución z_{S0}^2 de la ecuación en z^2.

- *Método W3 (el mejor método si todos los tamaños muestrales son menores o iguales que 10 y una buena alternativa en el resto de los casos)*

1) Incrementar todos los datos (los éxitos x_i y los fracasos y_i) en $z_{\alpha/2}^2 / 2K$ si $0 < x_i < n_i$ ($\forall i$) o, en otro caso, incrementarlos en

$$h_i = \begin{cases} \dfrac{z_{\alpha/2}^2}{2}.\dfrac{1+I_iK}{K} \text{ con } I_i = \begin{cases} 1 \text{ si } \bar{p}_i = \dfrac{1+s_i}{2} \\ 0 \text{ en otro caso} \end{cases} \text{ para } \lambda_I \\ \dfrac{z_{\alpha/2}^2}{2}.\dfrac{1+S_iK}{K} \text{ con } S_i = \begin{cases} 1 \text{ si } \bar{p}_i = \dfrac{1-s_i}{2} \\ 0 \text{ en otro caso} \end{cases} \text{ para } \lambda_S \end{cases} \quad \text{con } s_i = \begin{cases} +1 \text{ si } \beta_i > 0 \\ -1 \text{ si } \beta_i < 0 \end{cases}$$

2) El intervalo y el estadístico de contraste vienen dados, respectivamente, por las expresiones siguientes aplicadas a los datos incrementados anteriores:

$$L \in \bar{L} \pm z_{\alpha/2}\sqrt{\sum\frac{\beta_i^2\bar{p}_i\bar{q}_i}{n_i}}\ ,\ z_{W3}^2 = \left(\bar{L}-\lambda\right)^2 \Big/ \sum\frac{\beta_i^2\bar{p}_i\bar{q}_i}{n_i}$$

- *Método P0 (método muy conservador pero que no falla casi nunca)*

El intervalo y el estadístico de contraste vienen dados, respectivamente, por las expresiones siguientes (en las que $B=\Sigma\beta_i$):

$$L \in \frac{n}{n+z_{\alpha/2}^2}\left\{\bar{L} + \frac{Bz_{\alpha/2}^2}{2n} \pm \frac{z_{\alpha/2}}{2}\sqrt{\left(\sum\frac{\beta_i^2}{n_i}\right)\frac{n+z_{\alpha/2}^2}{n} - \left(\frac{B-2\bar{L}}{n}\right)^2}\right\},\ z_{P0}^2 = \frac{4\left(\bar{L}-\lambda\right)^2}{\sum\frac{\beta_i^2}{n_i} - \left(\frac{B-2\lambda}{n}\right)^2}$$

Como la aplicación del método de las marcas (con y sin cc) requiere de un proceso iterativo, Martín *et al.* [19, 20] desarrollaron un programa gratuito que permite la resolución del mismo para este tipo de inferencias. Dicho programa, que se encuentra en el enlace web siguiente: http://www.ugr.es/local/bioest/Z_LINEAR_K.EXE, proporciona los resultados según los procedimientos S0, S0c, W3 y P0, indicando en cada caso cual es el método óptimo a escoger.

6. EJEMPLOS

Price and Bonett [31] aluden a un estudio de Cohen *et al.* [9] en el que se anotó la presencia o ausencia de un tumor en cuatro grupos de 30 ratas sometidas a cuatro dietas (con mayor o menor grasa y con o sin fibra). La Tabla 2 muestra los resultados obtenidos y los tres contrastes de interés para evaluar el efecto de la fibra (L_2), para evaluar el efecto del nivel de grasa (L_3) o para evaluar la interacción entre ambos efectos (L_1: la diferencia entre los efectos de la fibra según que se determinen a uno u otro nivel de grasa). En todos los casos $\sum\beta_i = 0$ (por lo que se trata de contrastes).

	Fibra	*CON*		*SIN*	
	Grasa	*Alta*	*B*	*Al*	*B*
	Sample size (n_i)	30	30	30	30
(x_i)	*Rats showing cancer*	20	14	27	19
	Efectos	β_1	β_2	β_3	β_4
	L_1=Fibra×Grasa	+1	–1	–1	+1
	L_2=Fibra	+1	+1	–1	–1
	L_3=Grasa	+1	–1	+1	–1

Tabla 2 [31]

Tebbs and Roths [37] aluden a los datos (Tabla 3) de un ensayo clínico multicéntrico cuyo fin era evaluar la eficacia de un régimen rebajado en sal en el tratamiento de bebés varones con diarrea aguda. Una de las características a medir era el número de niños que experimentan fiebre con la administración del tratamiento o durante el ensayo. El objetivo es estimar el porcentaje de sujetos que responden al tratamiento. A causa de que el nivel de participación es diferente, pues depende de la localización, una estimación natural de la proporción global es la media de las probabilidades de respuesta p_i en los K=5 lugares, es decir $L=\sum \beta_i p_i$ con $\beta_i = n_i / \sum n_i$ (este modo de proceder es el adecuado si acaso los tamaños muestrales son proporcionales a los tamaños poblacionales). Ahora $\sum \beta_i \neq 0$.

Localización	*Tamaño de muestra (n_i)*	*Casos de fiebre (x_i)*	*Coeficientes (β_i)*
Bangladesh	158	73	158/675
Brasil	107	32	107/675
India	175	44	175/675
Perú	92	34	92/675
Vietnam	143	104	143/675
Total	675	287	1

Tabla 3 [37]

Si deseamos saber si el test de interacción para la Tabla 4 (H_0: L_1=0 *vs.* H_1: $L_1 \neq 0$) es significativo (a un error α=5%) sin realizar demasiados cálculos, podemos utilizar la regla de la expresión (5). Para el ejemplo se tiene que: λ=B=0, $z^2_{\alpha/2}=z^2_{2.5\%}$ =1.96^2 y $\bar{L}_1=-2/30, C=-15\times 1.96^2$, R_i^2 = 30×{30+1.96^2 (a_i+7.5×1.96^2)} con a_i=10, 2, –24 y 8 para i=1, 2, 3 y 4 respectivamente. Por ello, el valor de C en la función es $y(C)$=120 – $\sum R_i$ = –129.866 < 0, luego el test no es significativo.

El método S0 aplicado a los contrastes L_1, L_2 y L_3 proporciona los valores z_S = –0.412, –2.424 y +2.803 respectivamente, lo que indica que son significativos los efectos de la fibra (L_2) y de la grasa (L_3), pero que no hay interacción entre ambas (L_1) como se comprobó en el párrafo anterior. Para cuantificar la magnitud de los efectos de L_2 y L_3 hay que determinar el IC para cada una de ellas. Alternativamente, los tests anteriores pueden realizarse a través del IC para L_1, L_2 y L_3.

La Tabla 4 contiene los IC al 95% para todos los contrastes de la Tabla 2 y de la Tabla 3 realizados por los métodos S0, W3 y Pa0. Puede observarse que, como se indicó arriba, los contrastes L_2 y L_3 son significativos al error del 5% (pues sus IC contienen al valor 0), pero el contraste L_1 no lo es (pues sus IC no contienen al valor 0). Sin embargo, en la evaluación de la magnitud de los diversos intervalos para L_i o L se producen algunas diferencias entre métodos. Se observa que el método P0 proporciona unos IC excesivamente amplios, salvo para tamaños muestrales grandes (como en el caso de L). Finalmente se ve que el método W3 proporciona unos IC de anchura similar a la del método S0, pero sus centros son algo más diferentes (salvo en el caso de L, de nuevo por causa de los grandes valores de n_i).

IC (Tablas 4 y 5) = Centro (1ª entrada) ± Radio (2ª entrada)				
Método	L_1	L_2	L_3	L
0	–0,0719 ± 0,3164	–0,3934 ± 0,3162	0,4581 ± 0,3161	0,4256 ± 0,0349
3	–0,0646 ± 0,3162	–0,3876 ± 0,3162	0,4522 ± 0,3162	0,4256 ± 0,0348
0	–0,0646 ± 0,3520	–0,3876 ± 0,3454	0,4522 ± 0,3428	0,4256 ± 0,0372

Tabla 4: Análisis de los datos de las Tablas 2 y 3

La Tabla 5 muestra las salidas obtenidas al aplicar el programa Z_LINEAR_K.EXE [19, 20] para cada uno de los ejemplos planteados.

Z_LINEAR_K (2012) - Version 2.4

The program provides the asymptotical inferences (z-statistic or confidence intervals CI) for the parameters L=Beta1*p1 +...+ BetaK*pK (a lineal function of K independent proportions) or R=p2/p1 (the relative risk), by means of the optimal procedures and the score procedure.The key are:

xi=Binomial variable (ni=sample size; pi=probability)
TEST: H: L=lambda vs. K: L=/lambda or H: R=rho vs. K: R=/rho
CI (confidence interval): L1<=L<=L2 or R1<=R<=R2
cc=continuity correction

Test e intervalo de confianza para la interacción (L1)

Sample=	xi	ni	Betai	pi_LH
1	20	30	1.00	0.6850
2	14	30	-1.00	0.4457
3	27	30	-1.00	0.8918
4	19	30	1.00	0.6526
Totals	a1=80	n=120	B=0.00	B(-)=-2.00 B(+)=2.00

PARAMETER=L, ERROR(%)=5.00, LAMBDA=0.00, L(sample)=-0.0667

METHOD	$\lvert Z\rvert$-STATISTIC	P-VALUE	OPTIMAL IF
Score (without cc)	0.4119	0.6804	Some ni>10 and K>3
Score (with cc)	0.4118	0.6805	Some ni>10 and K=3
Adjusted Wald	0.4004	---	ni<=10 for all i (otherwise is a bit conservative)
Peskun	0.3651	0.7150	Much conservative, but almost never fails

METHOD	L1	L2	OPTIMAL IF
Score (without cc)	-0.3882	0.2445	Some ni>10 and K>3
Score (with cc)	-0.3882	0.2445	Some ni>10 and K=3
Adjusted Wald	-0.3808	0.2516	ni<=10 for all i (otherwise is a bit conservative)
Peskun	-0.4167	0.2875	Much conservative, but almost never fails

Test e intervalo de confianza para el efecto de la fibra (L2)

Sample=	xi	ni	Betai	pi_LH
1	20	30	1.00	0.7568
2	14	30	1.00	0.5855
3	27	30	-1.00	0.8313
4	19	30	-1.00	0.5110
Totals	a1=80	n=120	B=0.00	B(-)=-2.00 B(+)=2.00

PARAMETER=L, ERROR(%)=5.00, LAMBDA=0.00, L(sample)=-0.4000

METHOD	$\lvert Z\rvert$-STATISTIC	P-VALUE	OPTIMAL IF
Score (without cc)	2.4241	0.0153	Some ni>10 and K>3
Score (with cc)	2.4241	0.0153	Some ni>10 and K=3
Adjusted Wald	2.4023	---	ni<=10 for all i (otherwise is a bit conservative)
Peskun	2.1909	0.0285	Much conservative, but almost never fails

METHOD	L1	L2	OPTIMAL IF
Score (without cc)	-0.7096	-0.0772	Some ni>10 and K>3
Score (with cc)	-0.7096	-0.0772	Some ni>10 and K=3
Adjusted Wald	-0.7038	-0.0714	ni<=10 for all i (otherwise is a bit conservative)
Peskun	-0.7329	-0.0422	Much conservative, but almost never fails

Test e intervalo de confianza para el efecto del nivel de grasa (L3)

Sample=	xi	ni	Betai	pi_LH
1	20	30	1.00	0.5267
2	14	30	-1.00	0.6012
3	27	30	1.00	0.8156
4	19	30	-1.00	0.7411
Totals	a1=80	n=120	B=0.00	B(-)=-2.00 B(+)=2.00

PARAMETER=L, ERROR(%)=5.00, LAMBDA=0.00, L(sample)=0.4667

METHOD	$\lvert Z\rvert$-STATISTIC	P-VALUE	OPTIMAL IF
Score (without cc)	2.8033	0.0051	Some ni>10 and K>3
Score (with cc)	2.8033	0.0051	Some ni>10 and K=3
Adjusted Wald	2.8027	---	ni<=10 for all i (otherwise is a bit conservative)
Peskun	2.5560	0.0106	Much conservative, but almost never fails

METHOD	L1	L2	OPTIMAL IF
Score (without cc)	0.1420	0.7742	Some ni>10 and K>3
Score (with cc)	0.1420	0.7742	Some ni>10 and K=3
Adjusted Wald	0.1360	0.7684	ni<=10 for all i (otherwise is a bit conservative)
Peskun	0.1094	0.7950	Much conservative, but almost never fails

Intervalo de confianza para L				
Sample=	xi	ni	Betai	
1	73	158	0.23	
2	32	107	0.16	
3	44	175	0.26	
4	34	92	0.14	
5	104	143	0.21	
Totals	a1=287	n=675	B=1.00	B(-)=0.00 B(+)=1.00
PARAMETER=L, CONFIDENCE(%)=95.00, L(sample)=0.4252				
METHOD	L1	L2	OPTIMAL IF	
Score (without cc)	0.3907	0.4604	Some ni>10 and K>3	
Score (with cc)	0.3907	0.4604	Some ni>10 and K=3	
Adjusted Wald	0.3908	0.4604	ni<=10 for all i (otherwise is a bit conservative)	
Peskun	0.3884	0.4628	Much conservative, but almost never fails	

Tabla 5: Resultados con Z_LINEAR_K.EXE

7. CONCLUSIÓN

En el presente trabajo se ha hecho un resumen de la investigación realizada por Martín et al [19, 20] acerca de las inferencias asintóticas sobre una combinación de varias proporciones binomiales independientes.

El problema se aborda principalmente desde el punto de vista del método de las marcas (equivalente al método de la chi-cuadrado clásica), demostrándose la idoneidad del mismo frente al resto de procedimientos. Como la aplicación del método requiere de un proceso iterativo, el lector puede aplicar el programa gratuito que puede obtener en http://www.ugr.es/local/bioest/ Z_LINEAR_K.EXE.

El trabajo proporciona también una demostración teórica del resultado heurístico de que el intervalo de confianza al 95% de Wald mejora si a cada muestra se le añade $2/K$ éxitos y fallos. Adicionalmente, se generaliza esta regla, de modo que un IC sencillo y fiable (aunque conservador) consiste en aplicar el clásico el método de Wald de la expresión (2) añadiendo a cada muestra h_i' éxitos y fallos, cantidad que se reduce a $z_{\alpha/2}^2/2K$ cuando $0<x_i<n_i$ en todas las muestras (que es lo más habitual).

Además se define un nuevo procedimiento, basado en el método de Peskun [29], el cual puede ser una buena alternativa a las dos metodologías anteriores mencionadas por tener un buen comportamiento. Asimismo, se dota a los métodos más potentes de una cc, destacando la óptima actuación del método de las marcas con cc y se señala también, la importancia de que cualquier estadístico de test verifique las propiedades de convexidad espacial y paramétrica.

Agradecimientos / *Acknowledgements:* Este trabajo ha sido financiado por el Ministerio Español de Educación y Ciencia; proyecto número MTM2009-08886 (y cofinanciado por el Fondo Europeo de Desarrollo Regional)./*This research was supported by the Spanish Ministry of Education and Science; grant number MTM2009-08886 (co-financed by the European Regional Development Fund).*

REFERENCIAS

[1] AGRESTI, A. (2003). Dealing with discreteness: making ‘exact’ confidence intervals for proportions, differences of proportions, and odds ratios more exact. **Statistical Methods in Medical Research**, 12, 3-21.

[2] AGRESTI, A. & CAFFO, B. (2000). Simple and effective confidence intervals for proportions and difference of proportions result from adding two successes and two failures. **The American Statistician**, 54 (4), 280-288.

[3] AGRESTI, A. & COULL, B. A. (1998). Approximate Is Better than "Exact" for Interval Estimation of Binomial Proportions. **The American Statistician**, 52 (2), 119-126.

[4] AGRESTI, A.; BINI, M.; BERTACCINI, B. & RYU, E. (2008). Simultaneous confidence intervals for comparing binomial parameters. **Biometrics**, 64, 1270-1275.

[5] BARNARD, G.A. (1947). Significance tests for 2×2 tables. **Biometrika**, 34, 123-138.

[6] BROWN, L. & LI, X. (2005). Confidence intervals for two sample binomial distribution. **Journal of Statistical Planning and Inference**, 130, 359-375.

[7] BROWN, L.D., CAI, T.T. & DASGUPTA, A. (2001). Interval estimation for a binomial proportion. **Statistical Science**, 16, 101-133.

[8] CHAN, I.S.F. (2003). Proving non-inferiority or equivalence of two treatments with dichotomous endpoints using exact methods. **Statistical Methods in Medical Research**, 12, 37-58.

[9] COHEN, L. A.; KENDALL, M. E.; ZANG, E.; MESCHTER, C. & D. P. ROSE (1991). Modulation of N-Nitrosomethylurea-Induced Mammary Tumor Promotion by Dietary Fiber and Fat. **Journal of National Cancer Institution**, 83, 496-501.

[10] COX, D.R. (1970). The continuity correction. **Biometrika**, 57, 217-219.

[11] CRANS, G.G. & SHUSTER, J.J. (2008). How conservative is Fisher's exact test? A quantitative evaluation of the two-sample comparative binomial trial. **Statistics in Medicine**, 27, 3598–3611.

[12] DANN, R.S. & KOCH, G.G. (2005). Review and evaluation of methods for computing confidence intervals for the ratio of two proportions & considerations for non-inferiority clinical trials. **Journal of Biopharmaceutical Statistics**, 15, 85-107.

[13] GART, J.J. & NAM, J. (1988). Approximate interval estimation of the ratio of binomial parameters: A review and corrections for skewness. **Biometrics**, 44, 323-338.

[14] HABER, M. (1980). A comparison of some continuity corrections for the chi-squared test on 2×2 tables. **Journal of the American Statistical Association**, 75, 510-515.

[15] HERRANZ TEJEDOR, I. & 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, 969-981.

[16] KANG, S-H & CHEN, J.J. (2000). An approximate unconditional test of non-inferiority between two proportions. **Statistics in Medicine**, 19, 2089-2100.

[17] KOOPMAN, P.A.R. (1984). Confidence intervals for the ratio of two binomial proportions. **Biometrics**, 40, 513-517.

[18] LANG, J. B. (2008). Score and profile likelihood confidence intervals for contingency table parameters. **Statistics in Medicine**, 27, 5975-5990.

[19] MARTÍN, A., ÁLVAREZ, M. & HERRANZ, I. (2011). Inferences about a linear combination of proportions. **Statistical Methods in Medical Research**, 20, 369 – 387.

[20] MARTÍN, A., HERRANZ, I. & ÁLVAREZ, M. (2012). The optimal method to make inferences about a linear combination of proportions. **Journal of Statistical Computation and Simulation**, 82 , 123 – 135.

[21] MARTÍN ANDRÉS, A. & HERRANZ TEJEDOR, I. (2003). Unconditional confidence interval for the difference between two proportions. **Biometrical Journal**, 45, 426-436.

[22] MARTÍN ANDRÉS, A. & HERRANZ TEJEDOR, I. (2004). Exact unconditional non-classical tests on the difference of two proportions. **Computational Statistics and Data Analysis**, 45 (2), 373-388.

[23] MARTÍN ANDRÉS, A. & HERRANZ TEJEDOR, I. (2007). Propiedades del estadístico z en el contexto del test de equivalencia de dos proporciones. **XI Conferencia Española de Biometría y I Encuentro Iberoamericano de Biometría.** Salamanca, pgs. 79-80.

[24] MARTÍN ANDRÉS, A. & HERRANZ TEJEDOR, I. (2010). Asymptotic inferences about a linear combination of two proportions. **JP Journal of Biostatistics**, 4, 253-277.

[25] MEE, R.W. (1984). Confidence Bounds for the difference between two probabilities. **Biometrics**, 40 (4), 1175-1176.

[26] NAM, JUN-MO (1995). Confidence limits for the ratio of two binomial proportions based on likelihood scores: Non-iterative method. **Biometrical Journal**, 37 (3), 375-379.

[27] NEWCOMBE, R.G. (1998). Interval estimation for the difference between independent proportions: comparison of eleven methods. **Statistics in Medicine**, 17, 873-890.

[28] NEWCOMBE, R.G. (2001). Estimating the difference between differences: measurement of additive scale interaction for proportions. **Statistics in Medicine**, 20 (19), 2801-2994.

[29] PESKUN, P.H. (1993). A new confidence interval method based on the normal approximation for the difference of two binomial probabilities. **Journal of the American Statistical Association**, 88, 656-661.

[30] PHILLIPS, K.F. (2003). A new test of non-inferiority for anti-infective trials. **Statistics in Medicine**, 22, 201-212.

[31] PRICE, R. M. & BONETT, D. G. (2004). An improved confidence interval for a linear function of binomial proportions. **Computational Statistics & Data Analysis**, 45**,** 449-456.

[32] RÖHMEL, J. & MANSMANN, U. (1999a). Unconditional non-asymptotic one-sided tests for independent binomial proportions when the interest lies in showing non-inferiority and/or superiority. **Biometrical Journal**, 41, 149-170.

[33] RÖHMEL, J. & MANSMANN, U. (1999b). Exact tests of equivalence and efficacy with a non-zero lower bound for comparative studies by I.S.F. Chan (Letters to the Editor). **Statistics in Medicine**,18, 1734-1737.

[34] SANTNER, T.J.; PRADHAN, V.; SENCHAUDHURI, P.; METHA, C.R. & TAMHANE, A. (2007). Small-sample comparison of confidence intervals for the diference of two independent binomial proportions. **Computational Statistics & Data Analysis**, 51, 5791-5799.

[35] SCHAARSCHMIDT, F.; SILL, M. & HOTHORN, L.A. (2008). Approximate Simultaneous Confidence Intervals for Multiple Contrasts of Binomial Proportions. **Biometrical Journal**, 50 , 782-792.

[36] STERNE, T.E. (1954). Some remarks on confidence of fiducial limits. **Biometrika**, 41, 275-278.

[37] TEBBS, J. M. & ROTHS, S. A. (2008). New large-sample confidence intervals for a linear combination of binomial proportions. **Journal of Statistical Planning and Inference**, 138, 1884-1893.

[38] WILSON, E.B. (1927). Probable inference, the law of succession, and statistical inference. **Journal of the American Statistical Association**, 22, 209-212.

[39] ZOU, G.; HUANG, W. & ZHANG, X. (2009). A note on confidence interval estimation for a linear function of binomial proportions. **Computational Statistics & Data Analysis**, 53, 1080-1085.