

10 2.10.11

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS



LA COMPARACION DE PROPORCIONES:

UNA REVISION DE PRUEBAS DE SIGNIFICANCIA,
INTERVALOS DE CONFIANZA Y AJUSTES
POR ESTRATIFICACION

T E S I S

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O
P R E S E N T A

MA. DOLORES DONJUAN MORALES

MEXICO, D. F.

OCTUBRE DE 1981



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

INDICE

	Página
PROLOGO	I
1. INTRODUCCION Y RESUMEN	1
2. REVISION DE LOS ARGUMENTOS ESTADISTICOS AUXILIARES.	3
PARTE I. EL ANALISIS DE TABLAS DE 2 X 2	
3. LA TEORIA EXACTA	5
4. LA TEORIA ASINTOTICA	17
5. ESTIMACION PUNTUAL DEL COCIENTE DE MOMIOS	19
6. LIMITES DE CONFIANZA APROXIMADOS PARA COCIENTES DE MOMIOS.	24
7. UN EJEMPLO NUMERICO	28
8. EL ANALISIS DE PROPORCIONES ORDENAS APAREADAS	38
PARTE II. LA COMBINACION DE TABLAS DE 2 X 2	
9. MODELO LOGISTICO Y TEORIA EXACTA	41
10. TEORIA ASINTOTICA	46
11. ESTIMACION PUNTUAL DEL COCIENTE DE MOMIOS COMUN	48
12. LIMITES DE CONFIANZA APROXIMADOS PARA EL COCIENTE DE MOMIOS COMUN	53
13. PRUEBAS DEL MODELO LOGISTICO	56
14. ROBUSTES EN EFICIENCIA RELATIVA DEL MODELO LOGISTICO	57
15. EJEMPLOS NUMERICOS	60
16. METODOS RELACIONADOS Y APLICACIONES	102

ANEXO 1	105
ANEXO 2	109
BIBLIOGRAFIA	113

PROLOGO

En la investigación en áreas muy variadas como biología, medicina, sociología, psicología, ingeniería, comunicación, economía, etc., es frecuente que las variables de interés sean categóricas y numéricas. Para el análisis de variables numéricas hay una gran cantidad de textos que incluyen las técnicas más sofisticadas como regresión multivariada, análisis de varianza y covarianza. Sin embargo para el análisis de las variables categóricas hay pocos textos y algunos de ellos son muy profundos. Es por esto que un artículo de revisión de técnicas para comparar proporciones es muy importante para las personas interesadas en analizar variables categóricas.

En esta tesis se tomó el artículo de John J. Gart "The comparison of proportions: a review of significance test, confidence intervals and adjustments for stratification". Review of The International Statistical Institute. Volume 39:2, 1971, ya que presenta una buena revisión del tema de comparación de proporciones.

Lo que se pretende con este trabajo es aclarar diferentes puntos en los que se basa la teoría que se va desarrollando de manera que clarifique los argumentos de la comparación de proporciones. Así el lector al irse introduciendo en este artículo podrá tener

una comprensión mas rápida y mejor del material presentado. Los ejemplos han sido trabajados con detalle para facilitar su asimilación por parte de los estudiosos del tema.

De esta manera el lector encontrará una forma fácil, sencilla y rápida de introducirse en el tema, pero, si desea obtener información adicional respecto de los temas tratados, encontrará una amplia bibliografía en las referencias que se citan a lo largo de este estudio.

De esta forma se contribuye modestamente en la comprensión y difusión de los diferentes estudios sobre la comparación de proporciones.

En este trabajo solo se incluyen las dos primeras partes del artículo señalado, ya que como material de tesis se consideró suficiente, de cualquier forma se seguirá, trabajando la tercera parte que es la combinación de tablas de $2 \times t$.

Para el cálculo de el límite de confianza exacto para cociente de momios se utilizó un programa por computadora elaborado por Thomas [65], se estuvo probando, pero no todos los resultados se pudieron corroborar, ya que el límite superior del intervalo resultó diferente al que se obtiene en el artículo; además que no hay una forma rápida de ir comparando el programa. El programa se incluye en el anexo 2 y los intervalos de confianza quedan tal como aparecen en el artículo. Se continuará trabajando sobre dicho programa para su mayor ejecución.

El estimador puntual del cociente de momios común, en donde Goodman sugirió un segundo orden de perfeccionamiento sobre el estimador de Birch, el cual está denotado por $\hat{\psi}_{bg}$ no se calculó ya que se requiere de estudios más profundos, además de tiempo, pero se puede consultar en Goodman [36]. Algunos de los datos que se obtienen en el artículo difieren de los cálculos que se efectúan durante el trabajo, esto es debido a las aproximaciones que se hacen y son diferencias muy pequeñas, en general se corroboraron todos los datos presentados en dicho artículo.

LA COMPARACION DE PROPORCIONES:

UNA REVISION DE PRUEBAS DE SIGNIFICANCIA, INTERVALOS DE CONFIANZA Y AJUSTES POR ESTRATIFICACION,¹

POR: JOHN J. GART
Instituto Nacional de Cancer,
Bethesda, Maryland.

1. INTRODUCCION Y RESUMEN

Probablemente el problema cuantitativo más común en la investigación biológica involucra la comparación de poblaciones en relación a un atributo particular. En su forma más simple se reduce a un análisis de una tabla de 2×2 o a una prueba para la diferencia entre dos proporciones; en una forma más compleja puede involucrar la comparación de varias proporciones ajustadas por una variable auxiliar.

Esta disertación revisa la teoría y aplicación de análisis estadísticos apropiados para datos como estos. La teoría está enunciada en el esquema general de modelos logísticos formulados por Cox [16], [21] (ver también Rasch [60]). Este análisis usa la noción de condicionamiento de la distribución de las pruebas estadísticas sobre las esta-

1 Disertación invitada en la 7a. Conferencia Internacional de Biometría en Hannover, Alemania, Agosto 1970.

dísticas auxiliares, la cual implica, en el mejor de los casos, el llamado análisis de tablas de contingencia con "totales marginales fijos". También se consideran brevemente otras formulaciones.

Esta disertación esta dividida en tres partes: la primera trata las tablas 2×2 , la segunda con la combinación de tablas de 2×2 , y la última con la combinación de tablas de $2 \times t$. Las dos últimas partes están estrechamente relacionadas con el análisis de la interacción de alto orden en tablas de contingencia y el ajuste de tablas de contingencia de datos multivariados. Estas partes son tratadas tangencialmente aquí: la primera ha sido revisada completamente, recientemente por Plackett [58] y la segunda ha sido tratada por Bishop [7], las cuales contienen una bibliografía extensa. La excelente disertación por Birch [5,6] trata muchas de las cuestiones que nosotros discutimos aquí. Para referencias adicionales, ver Cox [21].

2. REVISION DE LOS ARGUMENTOS ESTADISTICOS AUXILIARES

Es útil revisar el argumento teórico, señalado por Fisher [24] y discutido más explícitamente por Cox [17], el cual trata con totales marginales fijos en tablas de contingencia. La teoría se funda en el concepto de estadística suficiente y estadística auxiliar. Considera el problema general, que consiste en la prueba de la hipótesis nula $H_0: \Theta_1 = \Theta_1^0, \Theta_2$ (no especificada) contra la alternativa $H_1: \Theta_1 \neq \Theta_1^0, \Theta_2$ (no especificada) donde Θ_1 y Θ_2 pueden ser escalares o vectores. Nosotros obtenemos un vector de observaciones $\{X\}$, el cual tiene una verosimilitud que se puede factorizar como sigue,

$$L(\Theta_1, \Theta_2; X) = \exp[\tau_1 \Theta_1 + \tau_2 \Theta_2] j(\Theta_1, \Theta_2) K(X).$$

Las funciones de la derecha son únicamente función de los argumentos mostrados y no de otras variables, τ_1 y τ_2 son funciones de X solamente. Bajo estas condiciones por el teorema de factorización de Fisher, τ_1 y τ_2 son estadísticas suficientes minimales para Θ_1 y Θ_2 respectivamente y, cuando se está haciendo inferencia acerca de Θ_1 , τ_2 es llamada estadística auxiliar. La teoría general sostiene que las inferen

cias relacionadas con $\Theta_1 = \Theta_1^0$, deben basarse en la distribución condicional de T_1 (usando letra mayúscula para indicar variable aleatoria), dado τ_2 , es decir,

$$m(\tau_1 | \tau_2; \Theta_1^0, \Theta_2) = q(\tau_1 | \tau_2; \Theta_1^0)$$

donde la distribución condicional, así señalada, no será una función de Θ_2 (ver Lehmann [45 p. 52]. Si Θ_1 y τ_1 son escalares, entonces la prueba es insesgada y uniformemente más potente, y está basada en la distribución condicional de T_1 [45 p.p. 134 f.f.]. Si T_1 es un vector, la prueba estadística es generalmente una función de las componentes de T_1 . Otras discusiones de esta teoría ha sido dada por Buehler [9], Suerdrup [64] y Fraser [26]. Lindley [47], revisando el último libro, expresa alguna sospecha "de algún argumento que como este, solo sirve en algunas situaciones..."

Sin embargo, afortunadamente el modelo logístico permite ubicar todo nuestro análisis de proporciones dentro de este esquema teórico.

PARTE I

PARTE 1. EL ANALISIS DE TABLAS DE 2 x 2

3. LA TEORIA EXACTA

Considere dos distribuciones binomiales independientes, con parámetros p_1 y p_2 y muestras de tamaño n_1 y n_2 . La verosimilitud es

$$L(p_1, p_2; x, y) = \binom{n_1}{x} \binom{n_2}{y} p_1^x (1 - p_1)^{n_1 - x} p_2^y (1 - p_2)^{n_2 - y} \quad (3.1)$$

Esta función de verosimilitud puede relacionarse fácilmente con la teoría mencionada, aunque no es esencial para el argumento hacerlo así [45 p.p. 143 f.f.], si reparametrizamos a una forma logística:

$$p_1 = \frac{e^{\beta + \lambda/2}}{1 + e^{\beta + \lambda/2}}, \quad p_2 = \frac{e^{\beta - \lambda/2}}{1 + e^{\beta - \lambda/2}} \quad (3.2)$$

Esto implica que la proporción o razón de momios es:

$$\psi = \frac{p_1 (1 - p_2)}{p_2 (1 - p_1)} = e^\lambda$$

Sustituyendo tenemos que:

$$\psi = \frac{\left[\frac{e^{\beta + \lambda/2}}{1 + e^{\beta + \lambda/2}} \right] \left[1 - \frac{e^{\beta - \lambda/2}}{1 + e^{\beta - \lambda/2}} \right]}{\left[\frac{e^{\beta - \lambda/2}}{1 + e^{\beta - \lambda/2}} \right] \left[1 - \frac{e^{\beta + \lambda/2}}{1 + e^{\beta + \lambda/2}} \right]} =$$

$$= \frac{\left[\frac{e^{\beta + \lambda/2}}{1 + e^{\beta + \lambda/2}} - \frac{e^{\beta + \lambda/2} e^{\beta - \lambda/2}}{[1 + e^{\beta + \lambda/2}][1 + e^{\beta - \lambda/2}]} \right]}{\left[\frac{e^{\beta - \lambda/2}}{1 + e^{\beta - \lambda/2}} - \frac{e^{\beta - \lambda/2} e^{\beta + \lambda/2}}{[1 + e^{\beta - \lambda/2}][1 + e^{\beta + \lambda/2}]} \right]} =$$

$$= \frac{\frac{e^{\beta + \lambda/2} [1 + e^{\beta - \lambda/2}] - e^{\beta + \lambda/2} e^{\beta - \lambda/2}}{[1 + e^{\beta + \lambda/2}][1 + e^{\beta - \lambda/2}]} - \frac{e^{\beta - \lambda/2} [1 + e^{\beta + \lambda/2}] - e^{\beta - \lambda/2} e^{\beta + \lambda/2}}{[1 + e^{\beta - \lambda/2}][1 + e^{\beta + \lambda/2}]}}{\frac{e^{\beta - \lambda/2} [1 + e^{\beta + \lambda/2}] - e^{\beta - \lambda/2} e^{\beta + \lambda/2}}{[1 + e^{\beta - \lambda/2}][1 + e^{\beta + \lambda/2}]}} =$$

$$= \frac{e^{\beta + \lambda/2} + e^{\beta + \lambda/2} e^{\beta - \lambda/2} - e^{\beta + \lambda/2} e^{\beta - \lambda/2}}{e^{\beta - \lambda/2} + e^{\beta - \lambda/2} e^{\beta + \lambda/2} - e^{\beta - \lambda/2} e^{\beta + \lambda/2}}$$

$$= \frac{e^{\beta + \lambda/2}}{e^{\beta - \lambda/2}}$$

$$= e^{\beta + \lambda/2 - [\beta - \lambda/2]}$$

$$= e^{\beta+\lambda/2 - \beta+\lambda/2}$$

$$= e^{2(\lambda/2)}$$

$$= e^{\lambda}$$

Por tanto

$$\psi = e^{\lambda} \tag{3.2.1}$$

o

$$\ln \psi = \lambda = \ln \left[\frac{p_1}{(1-p_1)} \right] - \ln \left[\frac{p_2}{(1-p_2)} \right]$$

esto es, λ es la diferencia en logits*. El parámetro λ está relacionado a la diferencia en proporciones teóricas, si $\lambda = 0$, $\psi = 1$ y $p_1 = p_2$; mientras que β está relacionada con la magnitud general de las p 's.

En esta notación la función de verosimilitud será:

$$L(p_1, p_2; x, y) = \binom{n_1}{x} \binom{n_2}{y} p_1^x (1-p_1)^{n_1-x} p_2^y (1-p_2)^{n_2-y}$$

$$L(\lambda, \beta; x, y) = \binom{n_1}{x} \binom{n_2}{y} \left[\frac{e^{\beta+\lambda/2}}{1 + e^{\beta+\lambda/2}} \right]^x \left[1 - \frac{e^{\beta+\lambda/2}}{1 + e^{\beta+\lambda/2}} \right]^{n_1-x} \left[\frac{e^{\beta-\lambda/2}}{1 + e^{\beta-\lambda/2}} \right]^y \left[1 - \frac{e^{\beta-\lambda/2}}{1 + e^{\beta-\lambda/2}} \right]^{n_2-y} =$$

* La transformación logits de una proporción P se define como $\text{Log } P/(1-P)$.

$$\begin{aligned}
 &= \binom{n_1}{x} \binom{n_2}{y} e^{(\beta+\lambda/2)x} e^{(\beta-\lambda/2)y} \left[\frac{1}{1+e^{\beta+\lambda/2}} \right]^x \left[\frac{1}{1+e^{\beta-\lambda/2}} \right]^y \\
 &\quad \left[\frac{1+e^{\beta+\lambda/2} - e^{\beta+\lambda/2}}{1+e^{\beta+\lambda/2}} \right]^{n_1-x} \left[\frac{1+e^{\beta-\lambda/2} - e^{\beta-\lambda/2}}{1+e^{\beta-\lambda/2}} \right]^{n_2-y} \\
 &= \binom{n_1}{x} \binom{n_2}{y} e^{\beta x + (\lambda/2)x + \beta y - (\lambda/2)y} \left[\frac{1}{1+e^{\beta+\lambda/2}} \right]^{n_1-x+x} \left[\frac{1}{1+e^{\beta-\lambda/2}} \right]^{n_2-y+y} \\
 &= \binom{n_1}{x} \binom{n_2}{y} e^{(x-y)\lambda/2 + (x+y)\beta} \left[\frac{1}{1+e^{\beta+\lambda/2}} \right]^{n_1} \left[\frac{1}{1+e^{\beta-\lambda/2}} \right]^{n_2}
 \end{aligned}$$

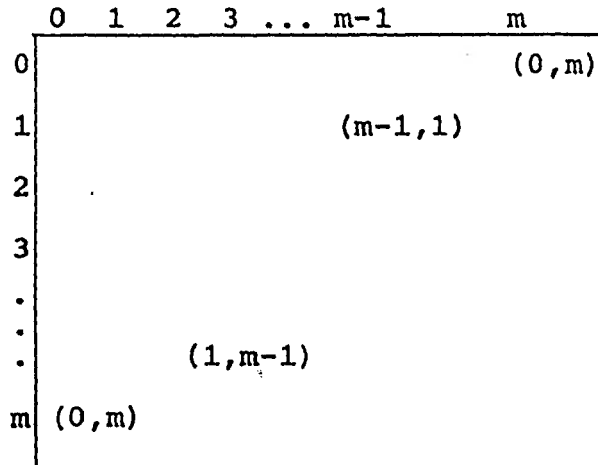
Por tanto:

$$L(\lambda, \beta; x, y) = \frac{\binom{n_1}{x} \binom{n_2}{y} e^{(x-y)\lambda/2 + (x+y)\beta}}{(1+e^{\beta+\lambda/2})^{n_1} (1+e^{\beta-\lambda/2})^{n_2}} \quad (3.3)$$

El conjunto de estadísticas suficientes es $x - y$ y $x + y$, las cuales son suficientes para λ y β respectivamente. Cuando hacemos inferencia sobre λ , entonces $x + y = m$ es la estadística auxiliar para el parámetro de estorbo β . Si encontramos la distribución condicional de $X - Y$ para m fijo, esto reduce la distribución condicional univariada de X la cual se encuentra fácilmente que es:

$$g(x|m; \lambda) = \frac{\binom{n_1}{x} \binom{n_2}{m-x} e^{\lambda x}}{\sum_{i=0}^m \binom{n_1}{i} \binom{n_2}{m-i} e^{\lambda i}} \quad (3.4)$$

Tenemos que:



en el denominador se suma sobre esta faja que es donde $x+y=m$, además sabemos que:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

de donde:

$$g(x|m, \lambda) = \frac{P(X=x, Y=m-x; \beta, \lambda)}{P(X+Y=m; \beta, \lambda)} = \frac{\binom{n_1}{x} \binom{n_2}{m-x} e^{[x-(m-x)]\lambda/2 + m\beta}}{(1+e^{\beta+\lambda/2})^{n_1} (1+e^{\beta-\lambda/2})^{n_2}} \cdot \frac{\sum_{i=0}^m \binom{n_1}{i} \binom{n_2}{m-i} e^{[i-(m-i)]\lambda/2 + m\beta}}{(1+e^{\beta+\lambda/2})^{n_1} (1+e^{\beta-\lambda/2})^{n_2}}$$

Por tanto

$$g(x|m, \lambda) = \frac{\binom{n_1}{x} \binom{n_2}{m-x} e^{\lambda x}}{\sum_{i=0}^m \binom{n_1}{i} \binom{n_2}{m-i} e^{\lambda i}}$$

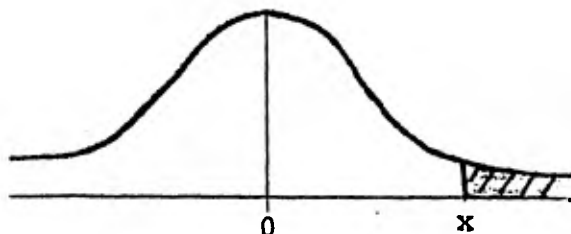
donde $x = 0, 1, \dots, m$. Este resultado fué, desde luego, señalado primero por Fisher [23]. Esta distribución puede usarse para diseñar pruebas de significancia o intervalos de confianza sobre algún valor de λ , o equivalentemente ψ .

La prueba de significancia de $H_0 : \lambda = \lambda_0$ contra $H_1 : \lambda > \lambda_0$ involucra el cálculo del valor P como sigue:

$$P = \sum_{i=x}^m g(i|m; \lambda_0) \quad (3.5)$$

donde P es la probabilidad de un extremo de la cola.

Se tienen aquí los valores de x 's grandes con una probabilidad P de ocurrir si $\lambda = \lambda_0$. Si P es muy pequeña, se rechaza la hipótesis $H_0 : \lambda = \lambda_0$ y se concluye $H_1 : \lambda > \lambda_0$.



Una secuencia de pruebas de una cola en ambas direcciones puede determinar un espacio hipotético de valores de $\lambda \in (\lambda_1, \lambda_2)$ que no entran en contradicción con los datos a un

nivel de significancia dado, digamos $\alpha/2$ en cada una de las colas. Esto es un intervalo de confianza con confianza $> 1 - \alpha$. Esto está formalmente definido por las ecuaciones siguientes [14,25];

$$\sum_{i=0}^x g(i|m; \lambda_2) = \alpha/2$$

$$\sum_{i=x}^m g(i|m; \lambda_1) = \alpha/2$$

o equivalentemente:

$$f_2(\psi_2) = \sum_{i=0}^x \binom{n_1}{i} \binom{n_2}{m-i} \psi_2^i / \sum_{i=0}^m \binom{n_1}{i} \binom{n_2}{m-i} \psi_2^i - \alpha_2 = 0$$

$$f_1(\psi_1) = \sum_{i=x}^m \binom{n_1}{i} \binom{n_2}{m-i} \psi_1^i / \sum_{i=0}^m \binom{n_1}{i} \binom{n_2}{m-i} \psi_1^i - \alpha_1 = 0$$

donde $\alpha_1 + \alpha_2 = \alpha$ y el coeficiente de confianza es $1 - \alpha$. Para obtener el valor de λ como límite de confianza inferior se toma $\log \psi_1$ y para el límite de confianza superior se toma $\log \psi_2$. Si $\lambda = 0$ (es decir $p_1 = p_2$) se contradice por los datos, entonces $\lambda = 0$ no estará contenido en el intervalo. Soluciones numéricas exactas para esos límites de confianza están disponibles usando un programa de computador obtenido por Thomas [65].

Desde luego esta cuestión se resuelve más fácilmente cuando $\lambda = 0$, considerando $g(x|m; \lambda = 0)$, el cual da la distribución hipergeométrica,

$$g(x|m; \lambda) = \frac{\binom{n_1}{x} \binom{n_2}{m-x} e^{\lambda x}}{\sum_{i=0}^m \binom{n_1}{i} \binom{n_2}{m-i} e^{\lambda i}}$$

$$g(x|m; \lambda = 0) = \frac{\binom{n_1}{x} \binom{n_2}{m-x} e^{0x}}{\sum_{i=0}^m \binom{n_1}{i} \binom{n_2}{m-i} e^0}$$

Como

$$\sum_{i=0}^m \binom{n_1}{i} \binom{n_2}{m-i} = \binom{n_1 + n_2}{m}$$

Entonces:

$$g(x|m; \lambda=0) = \frac{\binom{n_1}{x} \binom{n_2}{m-x}}{\binom{n_1 + n_2}{m}}, \quad x = 0, 1, 2, \dots, m$$

la prueba exacta de Fisher-Irwin está basada en (3.5).

Los momentos de la distribución condicional de X se puede escribir en términos de la función hipergeométrica [41] pero en el caso especial $\lambda = 0$, la media y varianza son más simples y quedan como sigue:

$$\begin{aligned} E(X|m; \lambda=0) &= \sum_{x=0}^m x g(x|m; \lambda=0) \\ &= \frac{\sum_{x=0}^m x \binom{n_1}{x} \binom{n_2}{m-x}}{\binom{n_1 + n_2}{m}} \end{aligned} \quad (3.5.1)$$

en esta expresión la x puede reducirse con otra x del denominador de $\binom{n_1}{x}$, con lo cual se obtiene:

$$\binom{n_1}{x} = \frac{n_1!}{x! (n_1 - x)!} = \frac{n_1 (n_1 - 1)!}{x(x-1)! (n_1 - 1 - x + 1)!}$$

$$x \binom{n_1}{x} = \frac{x n_1 (n_1 - 1)!}{x(x-1)! (n_1 - 1 - x + 1)!} = n_1 \binom{n_1 - 1}{x-1}$$

Por tanto

$$x \binom{n_1}{x} = n_1 \binom{n_1 - 1}{x-1}$$

y resulta

$$E(X|m; \lambda=0) = \frac{\sum_{x=1}^m n_1 \binom{n_1 - 1}{x-1} \binom{n_2}{m-x}}{\binom{n_1 + n_2}{m}}$$

en donde la suma se extiende de 1 a m , porque el primer término de (3.5.1) se anula y puede omitirse. Ahora bien: si en la última expresión se sustituye $x-1$ por y y se sacan los factores constantes del signo de suma, se obtiene;

$$E(X|m; \lambda=0) = \frac{n_1}{\binom{n_1 + n_2}{m}} \sum_{y=0}^{m-1} \binom{n_1 - 1}{y} \binom{n_2}{m-1-y}$$

Como:

$$\sum_{y=0}^{m-1} \binom{n_1 - 1}{y} \binom{n_2}{m-1-y} = \binom{n_1 - 1 + n_2}{m-1}$$

Entonces:

$$E(X|m; \lambda=0) = \frac{n_1}{\binom{n_1+n_2}{m}} \binom{n_1-1+n_2}{m-1}$$

Como

$$\begin{aligned} \binom{n_1+n_2}{m} &= \frac{(n_1+n_2)!}{m!(n_1+n_2-m)!} = \frac{(n_1+n_2)(n_1+n_2-1)!}{m(m-1)!(n_1+n_2-1-m+1)!} \\ &= \frac{n_1+n_2}{m} \binom{n_1+n_2-1}{m-1} \end{aligned}$$

Entonces

$$E(X|m; \lambda=0) = \frac{n_1 \binom{n_1+n_2-1}{m-1}}{\frac{n_1+n_2}{m} \binom{n_1+n_2-1}{m-1}} = \frac{n_1 m}{n_1+n_2}$$

Por tanto

$$E(X|m; \lambda=0) = \frac{n_1 m}{n_1+n_2}$$

Para calcular:

$V(X|m; \lambda=0)$ necesitamos el 2o. momento

$\sum_{x=0}^m x^2 g(x|m; \lambda=0)$ utilizaremos $x^2 = x(x-1) + x$ con

lo cual obtenemos

$$\sum_{x=0}^m x(x-1) g(x|m; \lambda=0) + \sum_{x=0}^m x g(x|m; \lambda=0)$$

hemos calculado la segunda suma al obtener la media, y aplicando el mismo procedimiento a la primera, se tiene:

$$\begin{aligned} \sum_{x=0}^m \frac{x(x-1) \binom{n_1}{x} \binom{n_2}{m-x}}{\binom{n_1+n_2}{m}} &= \frac{\sum_{x=2}^m n_1(n_1-1) \binom{n_1-2}{x-2} \binom{n_2}{m-x}}{\binom{n_1+n_2}{m}} \\ &= \frac{n_1(n_1-1)}{\binom{n_1+n_2}{m}} \sum_{y=0}^{m-2} \binom{n_1-2}{y} \binom{n_2}{m-2-y} \\ &= \frac{n_1(n_1-1)}{\binom{n_1+n_2}{m}} \binom{n_1+n_2-2}{m-2} = \frac{n_1(n_1-1) m(m-1)}{(n_1+n_2)(n_1+n_2-1)} \end{aligned}$$

Por tanto obtenemos el 2o. momento

$$\frac{n_1(n_1-1) m(m-1)}{(n_1+n_2)(n_1+n_2-1)} + \frac{n_1 m}{n_1+n_2}$$

Sabemos que $V(X) = E(X^2) - E(X)^2$, entonces:

$$\begin{aligned} V(X|m; \lambda=0) &= \frac{n_1(n_1-1) m(m-1)}{(n_1+n_2)(n_1+n_2-1)} + \frac{n_1 m}{n_1+n_2} - \left[\frac{n_1 m}{n_1+n_2} \right]^2 \\ &= \frac{n_1(n_1-1)m(m-1)(n_1+n_2) + n_1 m(n_1+n_2)(n_1+n_2-1) - (n_1 m)^2 (n_1+n_2+1)}{(n_1+n_2)(n_1+n_2-1)} \\ &= \left[\frac{1}{n_1+n_2(n_1+n_2-1)} \right] \left[n_1^3 m^2 - n_1^3 m - n_1^2 m^2 + n_1^2 m + n_1^2 n_2 m^2 - n_1^2 n_2 m \right] \end{aligned}$$

$$\begin{aligned}
 & - n_1 n_2 m^2 + n_1 n_2 m + n_1^3 m + n_1^2 n_2 m + n_1^2 n_2 m + n_1 n_2^2 m - n_1^2 m - n_1 n_2 m - n_1^3 m^2 - \\
 & - n_1^2 n_2 m^2 + n_1^2 m^2 \Big] \\
 & = \frac{1}{(n_1+n_2)^2 (n_1+n_2-1)} \left[-n_1 n_2 m^2 + n_1^2 n_2 m + n_1 n_2^2 m \right] = \frac{n_1 n_2 m (n_1 + n_2 - m)}{(n_1+n_2)^2 (n_1+n_2-1)}
 \end{aligned}$$

Por tanto

$$V(X|m; \lambda=0) = \frac{n_1 n_2 m (n_1 + n_2 - m)}{(n_1+n_2)^2 (n_1+n_2-1)}$$

Los momentos de la distribución hipergeométrica:

$$\left. \begin{aligned}
 E(X|m; \lambda=0) &= \frac{n_1 m}{n_1+n_2} \\
 V(X|m; \lambda=0) &= \frac{n_1 n_2 m (n_1 + n_2 - m)}{(n_1+n_2)^2 (n_1+n_2-1)}
 \end{aligned} \right\} (3.6)$$

4. LA TEORIA ASINTOTICA.

Si uno está interesado simplemente en probar la hipótesis nula, $\lambda=0$ para tablas de 2×2 y se dispone de números grandes, la estadística χ^2 (Ji-cuadrada) con un grado de libertad es apropiada:

$$\chi^2 = \frac{\{ |x - E(X|m; \lambda=0) | - 1/2 \}^2}{V(X|m; \lambda=0)} \quad (4.1)$$

o para pruebas de una cola la variable normal estandarizada $Z = \pm \sqrt{\chi^2}$, donde la corrección $1/2$ es aplicable solo porque estamos considerando todos los marginales como fijos.

Bajo la hipótesis alternativa, los resultados asintóticos están también disponibles. Stevens [63], Cornfield [14] y Hannan y Harkness [40] desarrollaron la siguiente teoría asintótica. Para números grandes la normalidad asintótica se obtiene con

$$\hat{E}(X|m; \lambda) = \hat{x} \quad (4.2)$$

y con

$$\hat{V}(X|m; \lambda) = \left[\frac{1}{\hat{x}} + \frac{1}{m - \hat{x}} + \frac{1}{n_1 - \hat{x}} + \frac{1}{n_2 - \hat{x} - m} \right]^{-1} \quad (4.3)$$

donde \hat{x} satisface la relación:

$$\psi = e^\lambda = \frac{\hat{x}(n_2 + \hat{x} - m)}{(n_1 - \hat{x})(m - \hat{x})} \quad (4.4)$$

Si encontramos el estimador de máxima verosimilitud de (3.1) estimando primero a p_1 y luego a p_2 , y sustituyendo en:

$$\hat{\psi} = \frac{\hat{p}_1(1 - \hat{p}_2)}{\hat{p}_2(1 - \hat{p}_1)}$$

tenemos que:

$$\hat{\psi} = \frac{\hat{x}(n_2 - \hat{x} - m)}{(m - \hat{x})(n_1 - \hat{x})} \quad \text{que resulta igual a (4.4)}$$

En la ecuación (4.4) hemos supuesto por simplicidad, que $m \leq n_1$. Para un valor de λ dado, o equivalentemente de ψ , la ecuación (4.4) conduce a una ecuación cuadrática en \hat{x} , donde solamente una raíz satisface $0 \leq \hat{x} \leq \min(n_1, m)$. Si $\lambda = 0$, es decir $\psi = 1$, entonces no es una ecuación cuadrática, pero los momentos exactos están disponibles en (3.6).

5. ESTIMACION PUNTUAL DEL COCIENTE DE MOMIOS.

El cociente de momios, ψ , o su logaritmo λ , es un parámetro clave en el análisis de las tablas de 2×2 y sus varias generalizaciones. Cornfield [14] demostró como puede usarse en el contexto de estudios retrospectivos y lo llamó riesgo relativo. Fisher [25] prefiere el término de cociente de productos cruzados. Edwardas [22] señala que ψ y funciones monótonas de ψ son únicas por poseer ciertas propiedades simétricas deseables.

De la estimación puntual de ψ surgen algunas preguntas no-triviales. Si maximizamos la verosimilitud incondicional (3.1), entonces encontramos el estimador de máxima verosimilitud a saber:

$$L(p_1, p_2; x, y) = \binom{n_1}{x} \binom{n_2}{y} p_1^x (1-p_1)^{n_1-x} p_2^y (1-p_2)^{n_2-y}$$

$$\ln L(p_1, p_2; x, y) = x \ln p_1 + (n_1 - x) \ln(1 - p_1) + y \ln p_2 +$$

$$+ (n_2 - y) \ln(1 - p_2) + \ln \left[\binom{n_1}{x} \binom{n_2}{y} \right]$$

$$\frac{\partial \ln L(p_1, p_2; x, y)}{\partial p_1} = \frac{x}{p_1} - \frac{(n_1 - x)}{1-p_1} = 0$$

$$x(1 - p_1) = p_1 (n_1 - x)$$

$$x - xp_1 = p_1 (n_1 - x)$$

$$x = p_1 (n_1 - x + x)$$

Por tanto

$$\hat{p}_1 = \frac{x}{n_1}$$

$$\frac{\partial \ln L(p_1, p_2; x, y)}{\partial p_2} = \frac{y}{p_2} - \frac{(n_2 - y)}{1 - p_2} = 0$$

Por tanto

$$\hat{p}_2 = \frac{y}{n_2}$$

Como

$$\hat{\psi} = \frac{\hat{p}_1 (1 - \hat{p}_2)}{\hat{p}_2 (1 - \hat{p}_1)}$$

Entonces

$$\hat{\psi} = \frac{x/n_1 (1 - y/n_2)}{y/n_2 (1 - x/n_1)}$$

$$= \frac{x(n_2 - y)/n_1 n_2}{y(n_1 - x)/n_1 n_2}$$

$$\hat{\psi}_{m1} = \frac{x(n_2 - y)}{y(n_1 - x)}$$

(5.1)

El estimador no es insesgado en la distribución incondicional de hecho su esperanza es infinita. Asintóticamente para $0 < p_1, p_2 < 1$, es un estimador insesgado de ψ con varianza [28].

$$v(\hat{\psi}) = \psi^2 \left[\frac{1}{n_1 p_1 (1-p_1)} + \frac{1}{n_2 p_2 (1-p_2)} \right] \quad (5.2)$$

En la escala logarítmica, la corrección de Haldane-Anscombe de 1/2 conduce al estimador aproximadamente insesgado

$$\hat{\psi} = \frac{\hat{p}_1 (1-\hat{p}_2)}{\hat{p}_2 (1-\hat{p}_1)}$$

$$\hat{\lambda} = \ln \hat{\psi} = \ln(\hat{p}_1/1-\hat{p}_1) - \ln(\hat{p}_2/1-\hat{p}_2)$$

$$= \ln \left[\frac{(x/n_1)/1-(x/n_1)}{n_1} \right] - \ln \left[\frac{(y/n_2)/1-(y/n_2)}{n_2} \right]$$

$$= \ln \left[\frac{\frac{x}{n_1}}{\frac{n_1-x}{n_1}} \right] - \ln \left[\frac{\frac{y}{n_2}}{\frac{n_2-y}{n_2}} \right]$$

$$= \ln \left[\frac{x}{n_1-x} \right] - \ln \left[\frac{y}{n_2-y} \right]$$

$$\hat{\lambda} = \ln \left[\frac{x+1/2}{n_1-x+1/2} \right] - \ln \left[\frac{y+1/2}{n_2-y+1/2} \right] \quad (5.3)$$

el cual tiene como varianza asintótica

$$\hat{V}(\hat{\lambda}) = \frac{1}{n_1 \hat{p}_1 (1 - \hat{p}_1)} + \frac{1}{n_2 \hat{p}_2 (1 - \hat{p}_2)},$$

la cual puede ser razonablemente bien estimada [33] mediante la siguiente expresión

$$\frac{1}{x + 1/2} + \frac{1}{n_1 - x + 1/2} + \frac{1}{y + 1/2} + \frac{1}{n_2 - y + 1/2} \quad (5.4)$$

El antilogaritmo de (5.3), $\hat{\psi} = e^{\hat{\lambda}}$, ha sido llamado "riesgo relativo corregido" por Lilienfeld [46]. Aunque no es "corregido" en el sentido usual de el término, tiene momentos finitos lo que no tiene el estimador no corregido de (5.1).

Los límites de confianza de $\ln \psi$ son aproximadamente:

$$\ln \psi_1 = \hat{\lambda} - X_\alpha \sqrt{\hat{V}(\hat{\lambda})}$$

$$\ln \psi_2 = \hat{\lambda} + X_\alpha \sqrt{\hat{V}(\hat{\lambda})}$$

Sacándoles antilogaritmos, obtenemos los límites de confianza correspondientes a ψ , son:

$$\psi_1 = e^{\hat{\lambda}} e^{-x_\alpha \sqrt{\hat{v}(\hat{\lambda})}}$$

$$\psi_2 = e^{\hat{\lambda}} e^{x_\alpha \sqrt{\hat{v}(\hat{\lambda})}}$$

El máximo de la verosimilitud condicional (3.4) no se encuentra fácilmente. Birch [5], usando resultados de Cox [16], ha probado que el estimador de máxima verosimilitud condicional es la solución apropiada, $\tilde{\psi}$, para la ecuación (ver también [36]),

$$x = E(x|m; \tilde{\psi}), \tag{5.5}$$

la cual es una ecuación polinomial de grados $\min(n_1, m)$. De la teoría asintótica, la moda de la verosimilitud asintótica debe ocurrir cuando $x = \hat{x}$, lo cual implica que la moda de la verosimilitud condicional asintótica es $\tilde{\psi}_{aml} = \hat{\psi}_{ml}$

6. LIMITES DE CONFIANZA APROXIMADOS PARA COCIENTES DE MOMIOS

Los resultados asintóticos del párrafo 4 pueden usarse para obtener límites de confianza para ψ (Cornfield [14], Fisher [25]). Para cualquier valor positivo de ψ , lo siguiente es una aproximación a la estadística ji-cuadrada.

$$\chi^2 = \frac{[\{x - E(X|m; \psi)\} \pm 1/2]^2}{\hat{V}(X|m; \psi)} \quad (6.1)$$

Fijando $\chi^2 = \chi_{\alpha}^2$, el porcentaje superior de la ji-cuadrada con un grado de libertad, tenemos dos ecuaciones cuárticas en \hat{x} , uno para $+ 1/2$ y otro para $-1/2$. La raíz más pequeña usando el signo menos se denota por \hat{x}_L , y la raíz más grande usando el signo más (+), se denota por \hat{x}_U . Los límites de confianza aproximados son:

$$\psi_L = \frac{\hat{x}_L (n_2 - m + \hat{x}_L)}{(m - \hat{x}_L) (n_1 - \hat{x}_L)} \quad (6.2)$$

y

$$\psi_U = \frac{\hat{x}_U (n_2 - m + \hat{x}_U)}{(m - \hat{x}_U) (n_1 - \hat{x}_U)} \quad (6.3)$$

Resolver las ecuaciones cuárticas no es un trabajo fácil, excepto con un programa de computación apropiado. Un

procedimiento iterativo relativamente fácil se puede usar con cualquier calculadora de escritorio, que calcule recíprocos fácilmente. El método involucrado en un pie de nota por Corntield [14], usa la relación iterativa,

$$\hat{x}_i = x \pm \frac{1}{2} \pm \chi\alpha \sqrt{\left[\frac{1}{\hat{x}_{i-1}} + \frac{1}{m-\hat{x}_{i-1}} + \frac{1}{n_1-\hat{x}_{i-1}} + \frac{1}{n_2+\hat{x}_{i-1}-m} \right]^{-1}}$$

(6.4)

donde $\hat{x}_0 = x$, y los signos positivos se aplican al límite superior y los signos negativos a el límite inferior. Cuando los números son pequeños es posible que \hat{x}_i pueda ser negativo o mayor que $\min(n_1, m)$. En este caso la técnica iterativa falla. En algunos casos las iteraciones proceden de acotar las soluciones sucesivas, e.g., $\hat{x}_i < \hat{x}_{i+1} < \hat{x}_{i-1}$, de manera que se acelera la convergencia por ir entre pares de valores sucesivos. Una técnica iterativa acelerada procede como sigue:

$$\hat{x}'_i = x \pm \frac{1}{2} \pm \chi\alpha \sqrt{\left[\frac{1}{\hat{x}^*_{i-1}} + \frac{1}{m-\hat{x}^*_{i-1}} + \frac{1}{n_1-\hat{x}^*_{i-1}} + \frac{1}{n_2+\hat{x}^*_{i-1}-m} \right]^{-1}}$$

(6.5)

$$\hat{x}_i^* = \frac{i \hat{x}_i' + (i-1) \hat{x}_{i-1}^*}{2i - 1},$$

donde $\hat{x}_0^* = x$. Este procedimiento convergerá algunas veces al rrededor del doble de rápido que el procedimiento previo. Si $x = 0$ solamente puede encontrarse un límite por lo que se sugiere hacer \hat{x}_0 o $\hat{x}_0^* = \frac{1}{2}$. Consideraciones similares se apli-
can cuando alguna de las otras entradas es cero.

Gart y Thomas [32] investigaron la exactitud a la solución de Cornfield por el método iterativo. Se encontró que el método de Cornfield es razonablemente preciso cuando la entrada más pequeña en la tabla de contingencia consistente con el límite excede 2, esto es, cuando $\min(\hat{x}, m-\hat{x}, n_1-\hat{x}, n_2-m+\hat{x}) \geq 2$. Muy a menudo cuando el método iterativo falla la solución no será exacta en cualquier caso. Algunas veces esto significa que el límite superior es exacto mientras que el límite infe-
rior no. También los límites al 95% pueden ser exactos mien-
tras los de 99% o mas puede no serlo.

Algunos otros métodos para encontrar límites de con-
fianza para cociente de momios han sido sugeridos también por Woolf [69], Gart [27], Goodman [34] y otros. Gart y Thomas [32] encontraron que el método logit, basado en (5.3) y (5.4) no es muy exacto, el intervalo es demasiado estrecho excepto para núme-
ros grandes. Para números pequeños una modificación muy precisa del méto-
do II de Gart [27] ha sido encontrada y será reportada en otra comunica-

ción [32], agregando reglas mas detalladas para escoger entre los métodos aproximados.

Debe notarse que los límites de confianza pueden también expresarse en términos de diferencias de proporciones

$$\hat{p}_1 - \hat{p}_2 = \frac{\hat{x}}{n_1} - \frac{m - \hat{x}}{n_2},$$

donde la \hat{x} se encuentra por la técnica iterativa o resolviendo la ecuación cuadrática (4.4) para el valor apropiado de ψ .

7. UN EJEMPLO NUMERICO.

La tabla 1 presenta los datos y el análisis de un experimento para probar el posible efecto carcinogénico de un insecticida Ayadex (2,3 - Dichloroallyl diisopropyl thiolcarbamate) en ratones, tal como fue reportado por Innes et al [42]. Los resultados son típicos de los métodos descritos en diversos aspectos. La P exacta y la dada por la prueba corregida de desviación normal estándar están muy cerca una de otra. El límite de confianza exacto al 95% incluye $\psi = 1$, como debe ser ya que la P exacta es mayor que 0.025. Los límites exactos y los de Cornfield son comparables, sin embargo, en la desviación relativa el límite inferior se desvía un poco más. Esto refleja el hecho de que el valor de $\hat{x} = 1.44$ usado para calcular el límite inferior es muy pequeño. Los límites logit son demasiado estrechos; de hecho entran en conflicto con la P aproximada, la cual excede 0.025, mientras que los límites logit no contienen $\psi = 1$. En el panel final de la tabla 1 las iteraciones sucesivas para la aproximación del límite inferior al 95% están dadas por las dos técnicas.

TABLA 1

Análisis de la "raza X" de ratones machos para frecuencias en exceso de tumores pulmonares. Cuando se alimentaron con el insecticida a 560 ppm por 85 semanas

(Fuente: Innes et al [42])

Ratones	Tratados	Control	Total
Con tumor	$x=4$	$m-x=y=5$	$x+y=m=9$
Sin tumor	$n_1-x=12$	$n_2-m+x=74$	$n_1+n_2-m=86$
Totales	$n_1=16$	$n_2=79$	$n_1+n_2=95$

Prueba de significancia: P exacta

$$P = \sum_{i=x}^m g(i|m; \lambda = 0)$$

$$= \sum_{i=x}^m \frac{\binom{n_1}{i} \binom{n_2}{m-i}}{\binom{n_1+n_2}{m}}$$

$$g(i = 4 | m = 9; \lambda = 0) = \frac{\binom{16}{4} \binom{79}{5}}{\binom{95}{9}} = 0.0349094$$

$$g(i = 5 | m = 9; \lambda = 0) = \frac{\binom{16}{5} \binom{79}{4}}{\binom{95}{9}} = 0.0055855$$

$$g(i = 6 | m = 9; \lambda = 0) = \frac{\binom{16}{6} \binom{79}{3}}{\binom{95}{9}} = 0.000538952$$

$$g(i = 7 | m = 9; \lambda = 0) = \frac{\binom{16}{7} \binom{79}{2}}{\binom{95}{9}} = 0.0000299973$$

$$g(i = 8 | m = 9; \lambda = 0) = \frac{\binom{16}{8} \binom{79}{1}}{\binom{95}{9}} = 0.000000865307$$

$$g(i = 9 | m = 9; \lambda = 0) = \frac{\binom{16}{9} \binom{79}{0}}{\binom{95}{9}} = 0.00000000973623$$

$$P = \sum_{i=4}^9 g(i | m; \lambda=0) = 0.0410647$$

∴ P = 0.0411

$$z = \frac{\{[x - \hat{E}(x|m; \lambda = 0)] - 1/2\}}{\sqrt{\hat{V}(x|m; \lambda = 0)}}$$

$$\begin{aligned}\hat{E}(x|m; \lambda=0) &= \frac{n_1 m}{n_1 + n_2} \\ &= \frac{16(9)}{95} = 1.516\end{aligned}$$

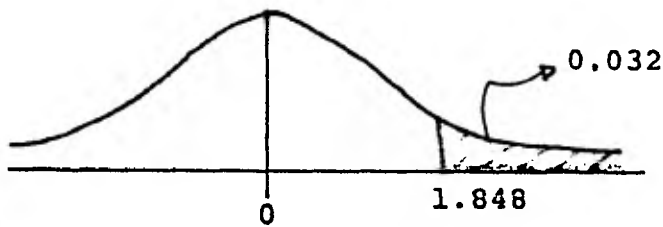
$$\begin{aligned}\hat{V}(x|m; \lambda = 0) &= \frac{n_1 n_2 m (n_1 + n_2 - m)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)} \\ &= \frac{(16)(79)(9)(95-9)}{(95)^2 (95-1)} \\ &= 1.153222137\end{aligned}$$

$$z = \frac{\{[4 - 1.516] - 1/2\}}{\sqrt{1.153222137}} = \frac{1.984}{1.073882} = 1.847503$$

∴ Z = 1.848

Buscando en la tabla de la Normal estandar $Z = 1.848$ tenemos que:

P aproximada = 0.032



Estimadores puntuales:

$$\begin{aligned}\hat{\psi}_{m1} &= \frac{x (n_2 - y)}{y (n_1 - x)} \\ &= \frac{4(79 - 5)}{5(16 - 4)} = 4.933333\end{aligned}$$

$$\therefore \boxed{\hat{\psi}_{m1} = 4.93}$$

$$\hat{\psi} = e^{\hat{\lambda}}; \hat{\lambda} = \ln \left[\frac{x + 1/2}{n_1 - x + 1/2} \right] - \ln \left[\frac{y + 1/2}{n_2 - y + 1/2} \right]$$

$$\hat{\lambda} = \ln \left[\frac{4 + 1/2}{16 - 4 + 1/2} \right] - \ln \left[\frac{5 + 1/2}{79 - 5 + 1/2} \right]$$

$$= \ln \left[\frac{4.5}{12.5} \right] - \ln \left[\frac{5.5}{74.5} \right]$$

$$= \ln(.36) - \ln(.073826)$$

$$= -1.021651 - (-2.606051)$$

$$= 1.5843997$$

$$e^{\hat{\lambda}} = e^{1.5843997} = 4.876363$$

$$\therefore \hat{\psi} = 4.88$$

$$\hat{p}_1 - \hat{p}_2 = \frac{x}{n_1} - \frac{m - x}{n_2}$$

$$= \frac{4}{16} - \frac{5}{79}$$

$$= 0.1867088$$

$$\therefore \hat{p}_1 - \hat{p}_2 = 0.187$$

Límites de Confianza al 95%:

Límites de Cornfield [14]

La solución iterativa para el límite inferior y superior al 95%, usando la iteración (6.4) y (6.5) se encuentra en el anexo 1.

$$\hat{x}_L = 1.44 \quad \text{para el límite inferior}$$

$$\hat{x}_U = 6.803 \quad \text{para el límite superior}$$

$$\psi_L = \frac{\hat{x}_L (n_2 - m + \hat{x}_L)}{(m - \hat{x}_L) (n_1 - \hat{x}_L)}$$

$$\psi_L = \frac{1.44(16 - 9 + 1.44)}{(9 - 1.44)(16 - 1.44)} = .9345892$$

$$\hat{p}_1 - \hat{p}_2 = \frac{\hat{x}_L}{n_1} - \frac{m - \hat{x}_L}{n_2}$$

$$\hat{p}_1 - \hat{p}_2 = \frac{1.44}{16} - \frac{9 - 1.44}{79} = - .0056962$$

$$\psi_U = \frac{\hat{x}_U (n_2 - m + \hat{x}_U)}{(m - \hat{x}_U) (n_1 - \hat{x}_U)}$$

$$\psi_U = \frac{6.803(79 - 9 + 6.803)}{(9 - 6.803)(16 - 6.803)} = 25.858445$$

$$\hat{p}_1 - \hat{p}_2 = \frac{\hat{x}_U}{n_1} - \frac{m - \hat{x}_U}{n_2}$$

$$\hat{p}_1 - \hat{p}_2 = \frac{6.803}{16} - \frac{9 - 6.803}{79} = .3973773$$

$$\therefore \boxed{.935 < \psi < 25.858}$$

$$\boxed{-.0057 < p_1 - p_2 < .3974}$$

Límites logit [27]

$$\psi_1 = e^{\hat{\lambda}} e^{-x_\alpha} \sqrt{\hat{V}(\hat{\lambda})} \text{ límite inferior}$$

$$\psi_2 = e^{\hat{\lambda}} e^{x_\alpha} \sqrt{\hat{V}(\hat{\lambda})} \text{ límite superior}$$

$$\hat{\lambda} = 1.5843997$$

$$e^{\hat{\lambda}} = 4.876363$$

$$\hat{V}(\hat{\lambda}) = \frac{1}{(16)(.25)(1-.25)} + \frac{1}{79(.0633)(1-.0633)}$$

$$\hat{V}(\hat{\lambda}) = .5468468$$

$$x_\alpha = x_{95} = 1.9596 \text{ ji-cuadrado al } 95\%$$

$$\psi_1 = 1.1458$$

$$\psi_2 = 19.5263$$

$$\therefore \boxed{1.16 < \psi < 19.52}$$

Para obtener los límites de $\hat{p}_1 - \hat{p}_2$ sustituimos

ψ_1 y ψ_2 en:

$$\psi = \frac{\hat{x}(n_2 - m + \hat{x})}{(n_1 - \hat{x})(m - \hat{x})}$$

Para ψ_1 :

$$1.16 = \frac{\hat{x}_U(79 - 9 + \hat{x}_U)}{(16 - \hat{x}_U)(9 - \hat{x}_U)}$$

$$.16 \hat{x}_U^2 - 99\hat{x}_U + 167.04 = 0$$

$$\hat{x}_U = \frac{99 \pm \sqrt{9801 - 106.9056}}{.32} = 1.71875$$

$$\therefore \hat{x}_U = 1.719$$

$$\hat{p}_1 - \hat{p}_2 = \frac{\hat{x}_U}{n_1} - \frac{m - \hat{x}_U}{n_2}$$

$$\hat{p}_1 - \hat{p}_2 = \frac{1.719}{16} - \frac{9 - 1.719}{79} = .0152729$$

Para ψ_2 :

$$19.52 = \frac{\hat{x}_L(79 - 9 + \hat{x}_L)}{(16 - \hat{x}_L)(9 - \hat{x}_L)}$$

$$18.52 \hat{x}_L^2 - 558 \hat{x}_L + 2810.88$$

$$\hat{x}_L = \frac{558 \pm \sqrt{311364 - 208229.99}}{37.04} = 6.3945777$$

$$\therefore \hat{x}_L = 6.395$$

$$\hat{p}_1 - \hat{p}_2 = \frac{\hat{x}_L}{n_1} - \frac{m - \hat{x}_L}{n_2}$$

$$\hat{p}_1 - \hat{p}_2 = \frac{6.395}{16} - \frac{9-6.395}{79} = .3667128$$

$$\therefore \boxed{.015 \leq \hat{p}_1 - \hat{p}_2 \leq .366}$$

Limites exactos [65] *

$$\boxed{0.834 \leq \psi \leq 26.16} , \quad \boxed{-0.0148 \leq p_1 - p_2 \leq 0.398}$$

* Consultar Thomas [65], el programa por computadora se encuentra en el anexo 2.

8. EL ANALISIS DE PROPORCIONES ORDENADAS APAREADAS.

Es un poco sorprendente que una situación experimental que involucre proporciones ordenadas apareadas conduzca a un análisis absolutamente equivalente al de la tabla general 2×2 [29]. Considere un ensayo clínico para comparar dos drogas A y B dadas secuencialmente a dos grupos de pacientes. A un grupo de n_a pacientes se les da la droga en orden (A,B) y sus respuestas se denotan por $Y_{ia} = 0$ ó 1 , y $Y_{ib} = 0$ ó 1 , $i = 1, 2, \dots, n_a$ dependiendo en que A y B respectivamente son efectivas o no. Similarmente a n_b pacientes se los da la droga en el orden (B,A) y las respuestas son denotadas por Y'_{ja} y Y'_{jb} , $j = 1, 2, \dots, n_b$. Deseamos comparar las drogas en relación a su efectividad. Si tratamos este problema dentro de un modelo logístico, debemos escribir:

$$P_r(Y_{ia}=1) = \frac{e^{\beta_i + \gamma + \lambda/2}}{1 + e^{\beta_i + \gamma + \lambda/2}},$$

$$P_r(Y_{ib}=1) = \frac{e^{\beta_i - \gamma - \lambda/2}}{1 + e^{\beta_i - \gamma - \lambda/2}}, \quad (8.1)$$

$$P_r(Y'_{ja}=1) = \frac{e^{\beta_i - \gamma + \lambda/2}}{1 + e^{\beta_i - \gamma + \lambda/2}},$$

$$P_r(Y'_{jb} = 1) = \frac{e^{\beta_i + \gamma - \lambda/2}}{1 + e^{\beta_i + \gamma - \lambda/2}} .$$

En analogía con la notación del análisis de varian-za β_i es el efecto del bloque, (o la pareja) γ es el efecto de orden y λ el efecto del tratamiento. Si suponemos que los resultados de los pacientes son mutuamente independientes, encontramos que la verosimilitud es:

$$\exp \left\{ \sum_{j=1}^{n_a} \beta_i (Y_{ia} + Y_{ib}) + \sum_{j=1}^{n_b} \beta_j (Y'_{ja} + Y'_{jb}) + \lambda (Y_{.a} + Y'_{.a} - Y_{.b} - Y'_{.b}) + \gamma (Y_{.a} + Y'_{.b} - Y_{.b} - Y'_{.a}) \right\}$$

$$\left\{ \prod_{i=1}^{n_a} (1 + e^{\beta_i + \lambda + \gamma}) (1 + e^{\beta_i - \lambda - \gamma}) \right\} \left\{ \prod_{j=1}^{n_b} (1 + e^{\beta_j - \lambda + \gamma}) (1 + e^{\beta_j + \lambda - \gamma}) \right\}$$

que es una función de las siguientes estadísticas suficientes conjuntamente

(i) $Y_{.a} + Y'_{.a} - Y_{.b} - Y'_{.b}$ la cual se relaciona a λ ,

(ii) $Y_{.a} + Y'_{.b} - Y_{.b} - Y'_{.a}$ la cual se relaciona a γ ,

(iii) $Y_{ia} = Y_{ib}$, $i=1,2,\dots,n_a$, la cual se relaciona a las β_i , y

(iv) $Y'_{ja} + Y'_{jb}$, $j=1,2,\dots,n_b$, la cual se relaciona a las β_j ,

donde la notación de punto es usada para indicar suma sobre el índice apropiado. La distribución condicional de (i) dando (ii), (iii) y (iv) fijos se simplifica mucho si notamos que la contribución de los pares, donde $Y_{ia} + Y_{ib}$ ó $Y'_{ja} + Y'_{jb} = 0$ ó 2, para (i), es cero. De este modo excluimos todos los pares semejantes y nos quedamos con n pares del orden (A,B) de los cuales Y_a tuvieron la droga efectiva A, y n' pares del orden (B,A) de los cuales Y'_a tuvieron la droga efectiva A. Con esta notación, encontramos que la distribución condicional de (i) dado (ii) es equivalente a Y_a dado $Y_a + Y'_a$. La distribución condicional relevante de este modo será:

$$h(Y_a | Y_a + Y'_a = m; \lambda) = \frac{\binom{n}{Y_a} \binom{n'}{m-Y_a} e^{\lambda Y_a}}{\sum_{i=0}^m \binom{n}{i} \binom{n'}{m-i} e^{\lambda i}}$$

la cual se puede ver que es enteramente equivalente a (3.4). Así todos los resultados subsecuentes en pruebas y límites de confianza se aplican a este problema como al análisis de la tabla general 2×2 . Por simetría se obtienen resultados similares para el efecto de orden γ . El lector se referirá a Gart [29] para detalles mas amplios.

PARTE II

PARTE II. LA COMBINACION DE TABLAS DE 2 x 2.

9. MODELO LOGISTICO Y TEORIA EXACTA.

Frecuentemente surge la cuestión de comparar dos tratamientos, por ejemplo con respecto a la proporción de respuesta, cuando los pacientes puedan estar estratificados o agrupados en bloques dentro de grupos por una covariable, digamos edad o sexo. Es bien conocido que simplemente sumar sobre los estratos puede ser engañoso en las tablas de 2 x 2 de resumen si los tamaños de muestra y la respuesta varía considerablemente en los grupos. Ejemplos obvios de esto han sido dados por Mantel y Haenszel [50], Gart [28], y Jowett [43].

Cochran [12] sugirió una prueba estadística aproximada para la combinación de tablas de 2 x 2. Se basa en tomar la media aritmética ponderada óptimamente de la diferencia en proporciones y suponer un modelo logístico. Mantel y Haenszel sugirieron una prueba estadística, la cual aparte de la corrección por continuidad y por población finita es equivalente a la estadística de Cochran. Su esquema de computación es también mucho más simple. Estos métodos son discutidos en dos libros de texto recientes y bien conocidos, [62, pp.253-257] y [8, pp.80-85].

La base teórica de la prueba estadística y los métodos relacionados establecidos elegantemente por Cox [19] en términos de la teoría de estadísticas auxiliares, está basada en un modelo logístico. Extendemos la notación de la sección 3 de una manera obvia. Considere b pares de variables binomiales independientes x_i y y_i , con parámetros correspondientes p_{i1} y p_{i2} , y tamaños de muestra n_{i1} y n_{i2} , donde $i = 1, 2, \dots, b$. La extensión natural del modelo logístico para una sola tabla de 2×2 supone:

$$p_{i1} = \frac{e^{\beta_i + \lambda/2}}{1 + e^{\beta_i + \lambda/2}}, \quad (9.1)$$

$$p_{i2} = \frac{e^{\beta_i - \lambda/2}}{1 + e^{\beta_i - \lambda/2}}, \quad i = 1, 2, \dots, b$$

donde las β_i 's son los efectos de bloque y la λ es el efecto del tratamiento. Es importante notar que el modelo logístico hace una importante suposición sobre las relaciones entre las Tablas de 2×2 ,

$$\log\left(\frac{p_{i1}}{1 - p_{i1}}\right) - \log\left(\frac{p_{i2}}{1 - p_{i2}}\right) = \lambda \quad (9.2)$$

esto es, que la diferencia en logits de las proporciones teóricas es constante en los diferentes estratos. Equivalentemente esto dice que el cociente de momios es constante, es decir:

$$\frac{p_{i1} (1 - p_{i2})}{p_{i2} (1 - p_{i1})} = \psi = e^\lambda. \quad (9.3)$$

NOTA: Esta igualdad se desarrolló en la parte I, (3.2.1) que es muy semejante a (9.3).

Esta suposición es equivalente a la no interacción de mayor orden en las tablas de contingencia de $2 \times 2 \times t$ en el sentido de Bartlett [2] y Roy y Kastenbaum [61].

Basado en este modelo la verosimilitud será:

$$L(\lambda, \beta_i; x_i, Y_i) = \frac{\exp\left\{\sum \left[(x_i + y_i)\beta_i \right] + (x. - y.)\lambda/2\right\}}{\prod_{i=1}^b (1 + e^{\beta_i + \lambda/2})^{n_{i1}} (1 + e^{\beta_i - \lambda/2})^{n_{i2}}}$$

$$y. = \sum_{i=1}^b y_i \quad x. = \sum_{i=1}^b x_i \quad (9.4)$$

NOTA: La verosimilitud se desarrolló en la parte I, (3.3) que es muy semejante a (9.4).

Vemos que el conjunto de estadísticas suficientes es $x_i + y_i$ para β_i , $i = 1, 2, \dots, b$ y $x. - y.$ para λ . Inferencias acerca de λ se deben basar en la distribución condicional de $x. - y.$ dado $x_i + y_i = m_i$ fijos. Para cada estrato tenemos que la distribución condicional de x_i es,

$$g_i(x_i | m_i; \lambda) = \frac{\binom{n_{i1}}{x_i} \binom{n_{i2}}{m_i - x_i} e^{\lambda x_i}}{\sum_{j=0}^{m_i} \binom{n_{i2}}{j} \binom{n_{i2}}{m_i - j} e^{\lambda j}} \quad (9.5)$$

y su distribución conjunta es

$$h(x_1, \dots, x_b | m_1, \dots, m_b; \lambda) = \prod_{i=1}^b g_i(x_i | m_i; \lambda) \quad (9.6)$$

Cuando las m_i son fijas la distribución de $X. - Y.$ es equivalente a la de $X.$, así las inferencias acerca de λ se deben hacer con la distribución condicional (9.6). Por ejemplo, la prueba de $H_0 : \lambda = \lambda_0$ contra $H_1 : \lambda > \lambda_0$ se puede probar mediante la computación de el valor exacto P ,

$$P = \sum_{j_i \in R} \prod_{i=1}^b g_i(j_i | m_i; \lambda_0) \quad (9.7)$$

donde R es el conjunto de todas las b -duplas particiones de toda $X. = \sum_i j_i$ tal que $X. \geq x.$, el valor observado de la estadística apropiada. De una manera similar fijando $P = \alpha/2$ en ambas direcciones se pueden encontrar los límites de confianza exactos al $1 - \alpha$ para λ y equivalentemente para ψ .

Si $\lambda_0 = 0$, es decir, $\psi = 1$, (9.5) se reduce a la distribución hipergeométrica y la P apropiada computada de (9.7), será la extensión natural de la prueba Fisher-Irwin dada por Cox [19] e implícita en Cochran [12]. Un programa de computa-

ción de esta prueba exacta está en preparación [68].

Los momentos de X . se expresan como una suma apropiada de los momentos en el caso simple. En particular, cuando $\lambda = 0$, tenemos:

$$\left. \begin{aligned} E(X | m_i; \lambda = 0) &= \sum_{i=1}^b \frac{n_{i1} m_i}{n_{i.}} \\ V(X | m_i; \lambda = 0) &= \sum_{i=1}^b \frac{n_{i1} n_{i2} m_i (n_{i.} - m_i)}{n_{i.}^2 (n_{i.} - 1)} \end{aligned} \right\} \quad (9.8)$$

NOTA: En la parte I se obtuvieron los momentos de la distribución hipergeométrica, (3.6) que son muy semejantes a (9.8).

10. TEORIA ASINTOTICA.

La teoría asintótica para el caso cuando $\lambda = 0$ es muy semejante al caso de una tabla simple 2×2 . Cuando n_{i1} y n_{i2} se hacen grandes, las X_i 's tienen distribuciones normales y de este modo X . está distribuída normalmente. Aún cuando algunas de las n_{i1} y n_{i2} sean pequeñas la X . tiene distribución normal para b grande por la forma de Liapounoff del teorema central del límite. La variable normal apropiada para la prueba de $\lambda = 0$, es:

$$Z = \frac{|x. - E(X. | m_i; \lambda = 0)| - 1/2}{\sqrt{V(X. | m_i; \lambda = 0)}} \quad (10.1)$$

Aparentemente las áreas de las colas dadas por esta aproximación normal, concuerdan muy bien con la prueba exacta (9.7). Por ejemplo, el problema de Cochran [12, pp.443 - 446] donde $b = 4$, la P exacta resultó 0.0095, mientras que de (10.1), $Z = 2.33$ y la P aproximada es 0.0098. En el ejemplo de Cox [19] donde $b = 18$, la P exacta fué 0.045, mientras que $Z = 1.64$ y la P aproximada 0.05.

Para $\psi \neq 1$ ó $\lambda \neq 0$, el valor asintótico de la me dia y varianza son: (recordando (4.2) y (4.3)),

$$\left. \begin{aligned} \hat{E}(X. | m_i; \lambda) &= \sum_{i=1}^b \hat{E}(X_i | m_i; \lambda), \\ \hat{V}(X. | m_i; \lambda) &= \sum_{i=1}^b \hat{V}(X_i | m_i; \lambda). \end{aligned} \right\} \quad (10.2)$$

donde, para poder conseguir resultados numéricos para un valor dado de λ o ψ , se tendrán que resolver b ecuaciones cuadráticas. Para aquellas tablas para la cual m_i es pequeña, será mejor y no muy difícil evaluar $E(X_i | m_i; \psi)$ exactamente comenzando con (9.5). Se puede suponer la distribución normal asintótica de X . para la X . estandarizada aunque aún está por estudiarse que tan adecuado resulta para muestras pequeñas.

11. ESTIMACION PUNTUAL DEL COCIENTE DE MOMIOS COMUN.

Los estimadores sugeridos para la ψ común pueden ser considerados de dos tipos, aquellos derivados o sugeridos por $x_i + y_i$ como aleatorios y aquellos por $x_i + y_i$ fijos. Para los casos marginales aleatorios, Norton [56] dió un esquema computacional para resolver las $t - 1$ ecuaciones de 3er grado involucradas en el método de estimación de máxima verosimilitud. El estimador se da en la columna cociente de su esquema. La varianza asintótica de este estimador es [28]:

para encontrar la varianza asintótica de

$$\hat{\psi}_{ML} = \frac{\hat{p}_{i1} (1 - \hat{p}_{i2})}{(1 - \hat{p}_{i1}) \hat{p}_{i2}}$$

reparametrizamos la verosimilitud eliminando b parámetros hechos redundantes por la restricción de la ψ común. Para hacer esto podemos hacer que:

$$\psi = \frac{p_{i1} (1 - p_{i2})}{(1 - p_{i1}) p_{i2}}$$

$$\psi (1 - p_{i1}) p_{i2} = p_{i1} (1 - p_{i2})$$

$$\psi p_{i2} - \psi p_{i1} p_{i2} = p_{i1} - p_{i1} p_{i2}$$

$$\psi p_{i2} - \psi p_{i1} p_{i2} - p_{i1} p_{i2} = p_{i1}$$

$$p_{i2}(\psi - \psi p_{i1} + p_{i1}) = p_{i1}$$

$$p_{i2}[\psi(1 - p_{i1}) + p_{i1}] = p_{i1}$$

$$\therefore p_{i2} = \frac{p_{i1}}{\psi(1 - p_{i1}) + p_{i1}}, \quad i = 1, 2, \dots, b$$

Procedemos a encontrar los elementos correspondientes a la matriz de información de los restantes $b + 1$ parámetros

$$E \left[\frac{-\partial^2 \ln L}{\partial \psi^2} \right] = \psi^{-2} \sum_{i=1}^b n_{i2} p_{i2} (1 - p_{i2}),$$

$$E \left[\frac{-\partial^2 \ln L}{\partial \psi \partial p_{i1}} \right] = -\psi^{-1} \frac{n_{i2} p_{i2} (1 - p_{i2})}{p_{i1} (1 - p_{i1})}, \quad i = 1, 2, \dots, b$$

$$E \left[\frac{-\partial^2 \ln L}{\partial p_{i1}^2} \right] = \frac{n_{i1} p_{i1} (1 - p_{i1}) + n_{i2} p_{i2} (1 - p_{i2})}{p_{i1}^2 (1 - p_{i1})^2}, \quad i = 1, 2, \dots, b$$

$$E \left[\frac{-\partial^2 \ln L}{\partial p_{i1} \partial p_{j1}} \right] = 0 \quad i \neq j.$$

Invirtiendo la matriz de información por el método habitual, encontramos la varianza asintótica de $\hat{\psi}_{ML}$

$$\therefore V(\hat{\psi}_{ML}) = \psi^2 w^{-1}$$

donde

$$w = \sum_{i=1}^b w_i$$

y

$$w_i = \left[\frac{1}{n_{i1} p_{i1} (1-p_{i1})} + \frac{1}{n_{i1} p_{i2} (1-p_{i2})} \right]$$

$$i = 1, 2, \dots, b$$

Woolf sugirió el estimador logit, el cual es esencialmente la medida geométrica ponderada, de los cocientes de medios individuales. Usando los factores de corrección sugeridos por Gart y Zweifel [33], podemos escribirlo como:

$$\hat{\psi}_g = \exp \left\{ \frac{\sum w_i \ln \left[\frac{(x_i + 1/2)(n_{2i} - y_i + 1/2)}{(y_i + 1/2)(n_{1i} - x_i + 1/2)} \right]}{\sum \hat{w}_i} \right\}, \quad (11.1)$$

donde

$$\hat{w}_i = \left[\frac{1}{x_i + 1/2} + \frac{1}{n_{i1} - x_i + 1/2} + \frac{1}{y_i + 1/2} + \frac{1}{n_{i2} - y_i + 1/2} \right]^{-1}, \quad (11.2)$$

para $i = 1, 2, \dots, b$. Gart [28] ha probado que este estimador es asintóticamente igualmente eficiente, que el estimador de máxima verosimilitud. Los otros dos estimadores, el de la media aritmética ponderada y la media armónica ponderada, discutidos por Gart, son también asintóticamente eficientes, pero carecen de propiedades deseables de invarianza que tiene $\hat{\psi}_g$

$$\hat{\psi}_a = \frac{\sum \hat{w}_i \hat{\psi}_i}{\sum \hat{w}_i}, \quad \text{estimador de la media aritmética ponderada.}$$

$$\hat{\psi}_h = \frac{\hat{w}}{\sum \hat{w}_i \hat{\psi}_i^{-1}}, \quad \text{estimador de la media armónica ponderada.}$$

Cox [21, pp.79-80] da una forma modificada de (11.1) la cual llamaremos $\hat{\psi}_{gc}$.

$$\hat{\psi}_{gc} = \frac{\sum \hat{w}_i \tilde{\Delta}_i^{(w)}}{\sum \hat{w}_i},$$

$\tilde{\Delta}_i^{(w)}$ Medida ponderada

$$\tilde{\Delta}_i^{(w)} \equiv \ln \left[\frac{x_i - 1/2}{n_{i1} - x_i - 1/2} \right] - \ln \left[\frac{y_i - 1/2}{n_{i2} - y_i - 1/2} \right]$$

Si consideramos la distribución condicional, o sea el caso de marginales fijos, fué probado por Birch [5] que el estimador de máxima verosimilitud de ψ es la solución de la ecuación polinomial

$$x = \sum_{i=1}^b E(X_i | m_i; \psi) \quad (11.3)$$

Esta es una ecuación de alto grado. Birch sugirió la siguiente solución aproximada, válida cuando $\hat{\psi} \sim 1$

$$\hat{\psi}_b = \exp \frac{\sum [x_i (n_{i2} - y_i) - y_i (n_{i1} - x_i)] / n_i}{\sum v(X_i | m_i) \lambda = 0}$$

Goodman [36, p.494] sugirió un segundo orden de perfeccionamiento (su $\exp(\theta^{**})$) sobre el estimador de Birch, el cual denotamos $\hat{\psi}_{bg}$. Otro estimador maximiza la verosimilitud asintótica, esto es, la solución $\hat{\psi}_{amI}$, a la ecuación

$$x = \sum_{i=1}^b \hat{E}(X_i | m_i; \psi), \quad (11.4)$$

la cual involucra la solución de b ecuaciones cuadráticas por ensayo y error. Un procedimiento iterativo para resolver esta ecuación se sugiere en la Tabla 3.

Mantel y Haenszel [50] sugirieron o revisaron 5 diferentes estimadores de ψ . Ellos prefieren el estimador,

$$\hat{\psi}_{mh} = \sum_{i=1}^b \left[\frac{x_i (n_{i2} - y_i)}{n_i} \right] / \sum_{i=1}^b \left[\frac{y_i (n_{i1} - x_i)}{n_i} \right]$$

el cual tiene la buena propiedad de que es igual a la unidad solamente cuando

$$x. = E (X. | m_i ; \psi = 1) .$$

12. LIMITES DE CONFIANZA APROXIMADOS PARA EL COCIENTE DE MOMIOS COMUN.

La teoría asintótica se puede usar para construir límites de confianza aproximados para la ψ común, los cuales son versiones extendidas de los límites de Cornfield para la tabla simple de 2×2 [31]. El límite superior se encuentra buscando el valor de ψ para el cual,

$$\frac{x. - \hat{E}(X. | m_i ; \psi) + 1/2}{\sqrt{\hat{V}(X. | m_i ; \psi)}} = - \chi_\alpha ,$$

y el límite inferior resolviendo

$$\frac{x. - \hat{E}(X. | m_i ; \psi) - 1/2}{\sqrt{\hat{V}(X. | m_i ; \psi)}} = + \chi_\alpha$$

(12.1)

En cada caso esto involucra la solución de b ecuaciones cuadráticas, tales como (4.4) y sustituir en (10.2). Esto involucra una solución de ensayo y error algo difícil. Un procedimiento iterativo para esta solución se sugiere en la Tabla 3.

Una primera aproximación para el límite puede obtenerse usando límites logit, sugerido primero por Woolf [69]. Estos son:

si consideramos $\ln \hat{\psi}_g$, en vez de $\hat{\psi}_g$, así también distribuido normalmente, hacemos primero para la ecuación estimada

$$\hat{w} (\ln \hat{\psi}_g - \ln \psi)^2 = \chi_\alpha^2$$

la solución de $\ln \psi$ es:

$$\hat{w}^{1/2} (\ln \hat{\psi}_g - \ln \psi) = \pm \chi_\alpha$$

$$\ln \hat{\psi}_g - \ln \psi = \pm \chi_\alpha \hat{w}^{-1/2}$$

$$\ln \psi = \ln \hat{\psi}_g \mp \chi_\alpha \hat{w}^{-1/2}$$

y tomando antilogaritmos, primero para el límite inferior

$$\psi_1 = \hat{\psi}_g \exp(-\chi_\alpha \hat{w}^{-1/2})$$

y luego para el límite superior

$$\psi_2 = \hat{\psi}_g \exp(\chi_\alpha \hat{w}^{-1/2})$$

(12.2)

Así con la evaluación de $\hat{\psi}_g$, ψ_1 y ψ_2 son más fácilmente computables con el uso de logaritmos.

Como en la tabla simple de 2 x 2, los límites logit parecen ser generalmente más angostos que los del método de Cornfield. Sin embargo, se usan como primeras aproximaciones en la resolución de (12.1). Los límites logit modificados pueden también estar basados en $\hat{\psi}_{gc}$ [21, pp.79-80]

$$\psi_1 = \hat{\psi}_{gc} \exp \{-\chi_\alpha \hat{W}^{-1/2}\}$$

$$\psi_2 = \hat{\psi}_{gc} \exp \{+\chi_\alpha \hat{W}^{-1/2}\}$$

Otros estimadores de intervalo pueden basarse en los estimadores puntuales de Birch y Goodman mencionados en la sección 11.

$$\psi_1 = \hat{\psi}_b \exp \left\{ -\chi_\alpha \left[\frac{1}{\sqrt{\sum V(X_i | m_i; \lambda = 0)}} \right] \right\}$$

$$\psi_2 = \hat{\psi}_b \exp \left\{ \chi_\alpha \left[\frac{1}{\sqrt{\sum V(X_i | m_i; \lambda = 0)}} \right] \right\}$$

Parece de el cálculo exacto de un ejemplo [31] que es preferible utilizar el método (12.1) a los otros.

Es también posible construir intervalos de confianza simultáneos aproximados sobre varias ψ 's individuales. Estos métodos han sido derivados por Cornfield [14] y Goodman [34].

13. PRUEBAS DEL MODELO LOGISTICO

La suposición del modelo logístico de que las ψ son constantes en los diferentes estratos es equivalente a la no existencia de la interacción de el orden más alto de las tablas de contingencia de $2 \times 2 \times b$ de Bartlett [2]. Esta suposición debe ser probada usando el procedimiento de cálculo de la Ji-cuadrada de la prueba de bondad de ajuste dada por Norton [56] o por el análisis logit sugerido por Woolf [69] (ver también [35]). En la notación de la sección 11, una versión modificada de la prueba logit es

$$X_{t-1}^2 = \sum_{i=1}^t \hat{\omega}_i (\ln \hat{\psi}_i - \ln \hat{\psi}_g)^2 \quad (13.1)$$

donde $\hat{\psi}_i$ son los cocientes de momios "corregidos" de la sección 5. El ajuste relativo de las estadísticas de prueba a la distribución Ji-cuadrada ha sido investigada numéricamente por Odoroff [57].

Zelen [70,71] ha obtenido pruebas exactas y aproximadas para la interacción sobre los espacios marginales fijos.

14. ROBUSTEZ EN EFICIENCIA RELATIVA DE EL MODELO LOGISTICO

El procedimiento para la combinación de tablas presentado arriba puede ser justificado por el modelo (9.1), si confiamos ya sea en el argumento de estadística suficiente o el equivalente, en este caso la teoría de pruebas insesgadas uniformemente más potentes. Como acabamos de ver, la suposición de λ constante puede ser probada, pero en las aplicaciones, esto nos puede conducir al problema familiar de pruebas estadísticas preliminares.

Ciertamente las tablas no deben ser combinadas a menos que exhiban una uniformidad aproximada de alguna clase.

Radhakrishna [59], siguiendo el enfoque de Cochran [12], consideró varios otros modelos de relación en las tablas. En general, puede suponerse que:

$$f(p_{i1}) - f(p_{i2}) = \Delta, \quad i = 1, 2, \dots, b,$$

donde $f(p)$ es una función monótona conocida para $0 < p < 1$, y Δ es una constante. Por ejemplo, para el modelo logístico $f(p) = \ln[p/(1-p)]$. La estadística para probar $\Delta = 0$ se supone que es de la forma de una combinación ponderada de la diferencia en proporciones.

$$\sum W_i (x_i/n_{i1} - y_i/n_{i2}), \quad (14.1)$$

y los pesos óptimos son escogidos de manera que se maximice la eficiencia relativa asintótica de la estadística de prueba para la elección particular de $f(p)$.

Radhakrishna ha derivado otras pruebas estadísticas para $f(p)$, como: logit, probit, raíz cuadrada de arcoseno, logarítmica, raíz cuadrada y funciones constantes. La prueba logit es, desde luego, equivalente a lo que se ha discutido. Él calculó la eficiencia relativa asintótica de cada estadística para situaciones donde las otras suposiciones se obtuvieron para una variedad de tamaños de muestra y variaciones en las p_i 's. Se encontró que los análisis logit, probit y arcoseno son muy similares en eficiencia sin importar cual de las tres es la suposición correcta. Sin embargo, sus eficiencias para las otras 3 suposiciones, particularmente la logarítmica fué tan baja como 62%. En general los análisis arcoseno son marginalmente mejores que el logit, sin embargo tiene el defecto que sus pesos (W_i en (14.1)) deben ser estimados, mientras que en el caso logit, son $n_{1i} n_{2i} / n_{.i}$ los cuales son, desde luego, constantes. De esto él concluye que el análisis logit es preferible con los tamaños de muestra moderados que son los que dispone en trabajos aplicados.

Otros análisis han sido sugeridos por Jowett [43], basado en una estandarización armónica de las tablas combinada 2×2 y Cochran [12] quien discutió una suma de las raíces cuadradas con signo de las Ji-cuadradas individuales de

las tablas 2 x 2. El último análisis está en algunas circunstancias, relacionado a el análisis arcoseno [59, p. 97]. Horbach [41a] sugirió todavía otras pruebas estadísticas y las comparó con las de Cochran y Radhakrisna teóricamente y por simulación en computadora para varias estructuras de probabilidad.

15. EJEMPLOS NUMERICOS

La tabla 2 da los datos extensos y sus análisis para el experimento de alimentación [42] referido previamente en la tabla 1. Los estratos son las cuatro categorías sexo-raza de las razas de ratones. El análisis estadístico individual de las tablas de 2×2 se presenta en la columna de la derecha. Solamente la primer tabla alcanzó un nivel de significancia del 5% (una-cola), sin embargo, por la prueba combinada la P aproximada (una-cola) es altamente significativa. La prueba para la interacción, por el método logit de (13.1), está lejos de la significancia, aún que uno puede estar inclinado a hacer pruebas de datos de hembras y machos separadamente. Los estimadores de la ψ común varían mucho, pero $\hat{\psi}_{mh}$ está muy cerca de $\hat{\psi}_{aml}$. El estimador de Birch, $\hat{\psi}_b$, se desvía apreciablemente, tal vez porque es bueno para ψ cercano a 1 y aquí ψ es muy grande. La versión corregida de Goodman, $\hat{\psi}_{bg}$, está un poco más cerca de $\hat{\psi}_{aml}$. Los límites de confianza para ψ calculados por cualquier método logit son otra vez bastante estrechos comparados con el método extendido de Cornfield. Se ve que el límite inferior al 99% de éste último concuerda con la P aproximada dado que $\alpha/2 = 0.005$, $\psi_L > 1.00$, y la P aproximada < 0.005 . Esta relación deseable entre estos dos métodos se mantendrá casi siempre. Ni la estimación por intervalos de Birch ni la de Goodman fueron cercanos a los límites extendidos de Cornfield para este ejemplo.

La tabla 3 da los detalles de un procedimiento iterativo para calcular los límites extendidos de Cornfield y $\hat{\psi}_{aml}$.

Algunos otros ejemplos fueron re-analizados y tienen varios aspectos que vale la pena notar. Birch [5], en su análisis de cultivo de tejidos de placenta por cesárea, ajustado por otros factores, encontró que el estimador exacto de máxima verosimilitud condicional de ψ es 2.66, mientras $\hat{\psi}_g = 2.36$, $\hat{\psi}_b = 2.55$, $\hat{\psi}_{aml} = 2.76$, y $\hat{\psi}_{mh} = 2.70$. La cercanía de $\hat{\psi}_{mh}$ y $\hat{\psi}_{aml}$ en estos dos ejemplos los señala como un buen punto inicial del estimador para solución por ensayo y error de uno de los estimadores de máxima verosimilitud.

Bliss [8, pp.82-83] analizó un ejemplo que involucra la mortalidad de pacientes del corazón con bloqueo de rama izquierda o derecha, ajustando por estratos de varios grupos de edad; por ambos métodos; el logit y el arcoseno. El primero produjo una Ji-cuadrada de 7.44 y el segundo una de 8.58. En este caso los dos análisis no dan una diferencia sustancial en interpretación. El ejemplo de Jowett [43] de diferencia en proporción de éxitos de 2 tratamientos ajustados por la categoría del paciente produjo una Ji-cuadrada $\chi^2 = 0.202$ por el método de estandarización armónica; el modelo logit da una prueba $\chi^2 = 0.438$. Estos ejemplos escogidos más bien aleatoriamente parecen confirmar que por lo menos para datos que se comportan regularmente, los diferentes métodos propuestos no presentan mucho desacuerdo entre sí.

TABLA 2.

Análisis de frecuencias de incremento de tumores pulmonares para 4 subgrupos de ratones alimentados con el fungicida Avadex [42]

SUBGRUPOS	TRATADOS	CONTROL	TOTALES	$\hat{\psi}$, P exacta y límites al 95%
X machos	con tumor $x_1 = 4$	$y_1 = 5$	$m_1 = 9$	$\hat{\psi}_{m1} = 4.93$
	sin tumor 12	74	86	P = 0,041
	Totales $n_{11} = 16$	$n_{12} = 79$	$n_{1.} = 95$	$0.834 \leq \psi \leq 26.16$
X hembras	con tumor $x_2 = 2$	$y_2 = 3$	$m_2 = 5$	$\hat{\psi}_{m1} = 4.00$
	sin tumor 14	84	98	P = 0.171
	Totales $n_{21} = 16$	$n_{22} = 87$	$n_{2.} = 103$	$0.303 \leq \psi \leq 9.36$
Y machos	con tumor $x_3 = 4$	$y_3 = 10$	$m_3 = 14$	$\hat{\psi}_{m1} = 2.29$
	sin tumor 14	80	94	P = 0.181
	Totales $n_{31} = 18$	$n_{32} = 90$	$n_{3.} = 108$	$0.455 \leq \psi \leq 9.36$
Y hembras	con tumor $x_4 = 1$	$y_4 = 3$	$m_4 = 4$	$\hat{\psi}_{m1} = 1.88$
	sin tumor 14	79	93	P = 0.495
	Totales $n_{41} = 15$	$n_{42} = 82$	$n_{4.} = 97$	$0.034 \leq \psi \leq 25.24$

Los valores de $\hat{\psi}_{m1}$ se obtuvieron como sigue:

$$\hat{\psi}_{m1} = \frac{\hat{P}_{i1} (1 - P_{i2})}{(1 - \hat{P}_{ii}) \hat{P}_{i2}}, \quad i = 1, \dots, 4$$

$$P_{i1} = \frac{x_i}{n_{i1}}, \quad P_{i2} = \frac{m_i - x_i}{n_{i2}}$$

$$\therefore \hat{\psi}_{m1} = \frac{x_i (n_{i2} - y_i)}{y_i (n_{i1} - x_i)}$$

Para X machos.

$$\hat{\psi}_{m1} = \frac{4(79 - 5)}{5(16 - 4)} = \boxed{4.93}$$

Para Y machos:

$$\hat{\psi}_{m1} = \frac{4(90 - 10)}{10(18 - 4)} = \boxed{2.2857}$$

Para X hembras:

$$\hat{\psi}_{m1} = \frac{2(87 - 3)}{3(16 - 2)} = \boxed{4.00}$$

Para Y hembras:

$$\hat{\psi}_{m1} = \frac{1(82 - 3)}{3(15 - 1)} = \boxed{1.88095}$$

P exacta

$$P = \sum_{i=x}^m g(i|m; \lambda = 0) = \frac{\sum_{i=x}^m \binom{n_1}{i} \binom{n_2}{m-i}}{\binom{n_1 + n_2}{m}}$$

Para X hembras:

$$P = \sum_{i=2}^5 g(i|m; \lambda = 0)$$

$$g(i = 2 | m = 5; \lambda = 0) = \frac{\binom{16}{2} \binom{87}{3}}{\binom{103}{5}} = .145296$$

$$g(i = 3 | m = 5; \lambda = 0) = \frac{\binom{16}{3} \binom{87}{2}}{\binom{103}{5}} = .0239311$$

$$g(i = 4 | m = 5; \lambda = 0) = \frac{\binom{16}{4} \binom{87}{1}}{\binom{103}{5}} = .00180874$$

$$g(i = 5 | m = 5; \lambda = 0) = \frac{\binom{16}{5} \binom{87}{0}}{\binom{103}{5}} = .0000498964$$

$$P = .1710857$$

∴ P = .171

Para Y machos:

$$P = \sum_{i=4}^{14} (i|m; \lambda = 0)$$

$$g(i = 4|m = 14; \lambda = 0) = \frac{\binom{18}{4} \binom{90}{10}}{\binom{108}{14}} = .1252713$$

$$g(i = 5|m = 14; \lambda = 0) = \frac{\binom{18}{5} \binom{90}{9}}{\binom{108}{14}} = .0433036$$

$$g(i = 6|m = 14; \lambda = 0) = \frac{\binom{18}{6} \binom{90}{8}}{\binom{108}{14}} = .0102978$$

$$g(i = 7|m = 14; \lambda = 0) = \frac{\binom{18}{7} \binom{90}{7}}{\binom{108}{14}} = .00170153$$

$$g(i = 8|m = 14; \lambda = 0) = \frac{\binom{18}{8} \binom{90}{6}}{\binom{108}{14}} = .000194967$$

$$g(i = 9|m = 14; \lambda = 0) = \frac{\binom{18}{9} \binom{90}{5}}{\binom{108}{14}} = .0000152915$$

$$g(i = 10|m = 14; \lambda = 0) = \frac{\binom{18}{10} \binom{90}{4}}{\binom{108}{14}} = 8.00139 \times 10^{-7}$$

$$g(i = 11|m = 14; \lambda = 0) = \frac{\binom{18}{11} \binom{90}{3}}{\binom{108}{14}} = 2.67549 \times 10^{-8}$$

$$g(i = 12 | m = 14; \lambda = 0) = \frac{\binom{18}{12} \binom{90}{2}}{\binom{108}{14}} = 5.32058 \times 10^{-10}$$

$$g(i = 13 | m = 14; \lambda = 0) = \frac{\binom{18}{13} \binom{90}{1}}{\binom{108}{14}} = 5.51832 \times 10^{-12}$$

$$g(i = 14 | m = 14; \lambda = 0) = \frac{\binom{18}{14} \binom{90}{0}}{\binom{108}{14}} = 2.18981 \times 10^{-14}$$

$$P = .1807853$$

∴

$P = .181$

Para Y hembras:

$$P = \sum_{i=1}^4 g(i|m; \lambda = 0)$$

$$g(i = 1 | m = 4 ; \lambda = 0) = \frac{\binom{15}{1} \binom{82}{3}}{\binom{97}{4}} = .3833943$$

$$g(i = 2 | m = 4 ; \lambda = 0) = \frac{\binom{15}{2} \binom{82}{2}}{\binom{97}{4}} = .100641$$

$$g(i = 3 | m = 4 ; \lambda = 0) = \frac{\binom{15}{3} \binom{82}{1}}{\binom{97}{4}} = .0107681$$

$$g(i = 4 | m = 4 ; \lambda = 0) = \frac{\binom{15}{4} \binom{82}{0}}{\binom{97}{4}} = .000393957$$

$$P = .4951973$$

$$\therefore \boxed{P = .495}$$

Nota: La $\boxed{P \text{ exacta} = 0.041}$ Para X machos se calculó en la tabla 1.

Para el cálculo de los intervalos de confianza (ψ) consultar Thomas [35], el programa por computadora se encuentra en el anexo 2.

Resumen de estadísticas

$$x. = \sum_{i=1}^4 x_i = 4 + 2 + 4 + 1 = 11 \quad \therefore \boxed{x. = 11}$$

$$E(X.) = E(X. | m_i; \lambda = 0) = \sum_{i=1}^4 \frac{n_{i1} m_i}{n_{i.}}$$

$$E(X.) = \frac{16 \cdot 9}{95} + \frac{16 \cdot 5}{103} + \frac{18 \cdot 14}{108} + \frac{15 \cdot 4}{97} = 5.2443785$$

$$\therefore \boxed{E(X.) = 5.24}$$

$$z = \frac{|x. - E(X. | m_i, \lambda = 0)| - 1/2}{\sqrt{V(X. | m_i; \lambda = 0)}}$$

$$V(X. | m_i; \lambda = 0) = \sum_{i=1}^4 \frac{n_{i1} n_{i2} m_i (n_{i.} - m_i)}{n_{i.}^2 (n_{i.} - 1)}$$

$$V(X. | m_i; \lambda = 0) = \frac{16 \cdot 79 \cdot 9 (95-9)}{95^2 (95-1)} + \frac{16 \cdot 87 \cdot 5 (103-5)}{103^2 (103-1)} +$$

$$\frac{18 \cdot 90 \cdot 14 (108-14)}{108^2 (108-1)} + \frac{15 \cdot 82 \cdot 4 (97-4)}{97^2 (97-1)} =$$

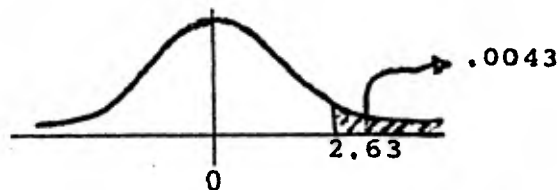
$$= 3.9983079$$

$$\therefore \boxed{V(X. | m_i; \lambda = 0) = 3.998}$$

$$z = \frac{|11 - 5.24| - 0.5}{1.9995769} = 2.6283668$$

∴ $z = 2.63$

Buscando en la tabla de la Normal estándar



$z = 2.63$ Tenemos que:

∴ $P \text{ aproximada} = .0043$

Prueba de interacción por el método logit

$$\chi^2_{t-1} = \sum_{i=1}^t \hat{w}_i (\ln \hat{\psi}_i - \ln \hat{\psi}_g)^2$$

$$\ln \hat{\psi}_i = \frac{(x_i + 1/2)(n_{i2} - y_i + 1/2)}{(y_i + 1/2)(n_{i1} - x_i + 1/2)}$$

$$\ln \hat{\psi}_g = \frac{\sum \hat{w}_i \ln \hat{\psi}_i}{\sum \hat{w}_i}$$

$$\hat{w}_i = \left[\frac{1}{x_i + 1/2} + \frac{1}{n_{i1} - x_i + 1/2} + \frac{1}{y_i + 1/2} + \frac{1}{n_{i2} - y_i + 1/2} \right]^{-1}$$

$$\ln \hat{\psi}_1 = \frac{(4+0.5)(79-5+0.5)}{(5+0.5)(16-4+0.5)} = \ln 4.8763636 = 1.5843999$$

$$\ln \hat{\psi}_2 = \frac{(2+0.5)(87-3+0.5)}{(3+0.5)(16-2+0.5)} = \ln 4.1625616 = 1.4261306$$

$$\ln \hat{\psi}_3 = \frac{(4+0.5)(90-10+0.5)}{(10+0.5)(18-4+0.5)} = \ln 2.3793103 = .8668106$$

$$\ln \hat{\psi}_4 = \frac{(1+0.5)(82-3+0.5)}{(3+0.5)(15-1+0.5)} = \ln 2.3497537 = .8543105$$

$$\begin{aligned} \hat{w}_1 &= \left[\frac{1}{4+0.5} + \frac{1}{16-4+0.5} + \frac{1}{5+0.5} + \frac{1}{79-5+0.5} \right]^{-1} \\ &= (.4974632)^{-1} = 2.0101989 \end{aligned}$$

$$\begin{aligned} \hat{w}_2 &= \left[\frac{1}{2+0.5} + \frac{1}{16-2+0.5} + \frac{1}{3+0.5} + \frac{1}{87-3+0.5} \right]^{-1} \\ &= (.7665141)^{-1} = 1.3046074 \end{aligned}$$

$$\begin{aligned} \hat{w}_3 &= \left[\frac{1}{4+0.5} + \frac{1}{18-4+0.5} + \frac{1}{10+0.5} + \frac{1}{90-10+0.5} \right]^{-1} \\ &= (.3988481)^{-1} = 2.5072196 \end{aligned}$$

$$\begin{aligned} \hat{w}_4 &= \left[\frac{1}{1+0.5} + \frac{1}{15-1+0.5} + \frac{1}{3+0.5} + \frac{1}{82-3+0.5} \right]^{-1} \\ &= (1.0339251)^{-1} = .967188 \end{aligned}$$

$$\sum_{i=1}^4 \hat{w}_i = 6.789214$$

$$\ln \hat{\psi}_g = \frac{\sum_{i=1}^4 \hat{w}_i \ln \hat{\psi}_i}{\sum \hat{w}_i}$$

$$\begin{aligned} \sum_{i=1}^4 \hat{w}_i \ln \hat{\psi}_i &= (2.0101989)(1.5843998) \\ &+ (1.3046074)(1.4261306) \\ &+ (2.5072196)(.8668106) \\ &+ (.967188)(.8543105) \\ &= 8.0450627 \end{aligned}$$

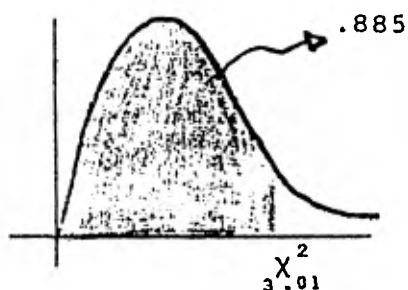
$$\ln \hat{\psi}_g = \frac{8.0450627}{6.789214} = 1.184977$$

$$\chi_{4-1}^2 = \sum_{i=1}^4 \hat{w}_i (\ln \hat{\psi}_i - \ln \hat{\psi}_g)^2$$

$$\begin{aligned} \chi_3^2 &= (2.0101989)(1.5843998-1.184977)^2 \\ &+ (1.3046074)(1.4261306-1.184977)^2 \\ &+ (2.5072196)(.8668106-1.184977)^2 \\ &+ (.967188)(.8543105-1.184977)^2 \\ &= .7561319 \end{aligned}$$

$$\therefore \chi_3^2 = .7561$$

Buscando en la tabla de Ji cuadrada



$$P = 1 - .115 = .885$$

$$\therefore P = .885$$

Estimadores puntuales:

$$\hat{\psi}_g = e^{1.184977} = 3.2706117$$

$$\therefore \boxed{\hat{\psi}_g = 3.27}$$

$$\hat{\psi}_b = \exp \left\{ \frac{\sum_{i=1}^4 [x_i (n_{i2} - y_i) - y_i (n_{i1} - x_i)] / n_i}{\sum V(X_i | m_i; \lambda = 0)} \right\}$$

$$V(X_i | m_i; \lambda = 0) = 3.9983079$$

$$\begin{aligned} \frac{\sum_{i=1}^4 [x_i (n_{i2} - y_i) - y_i (n_{i1} - x_i)]}{n_i} &= \frac{4(79 - 5) - 5(16 - 4)}{95} \\ &+ \frac{2(87-3) - 3(16-2)}{95} \\ &+ \frac{4(90-10) - 10(18-4)}{108} \\ &+ \frac{1(82-3) - 3(15-1)}{97} = \end{aligned}$$

$$= 5.7556215$$

$$\hat{\psi}_b = \exp \left\{ \frac{5.7556215}{3.9983079} \right\} = e^{1.4395143} = 4.2186464$$

$$\therefore \hat{\psi}_b = 4.22$$

$$\hat{\psi}_{mh} = \frac{\sum_{i=1}^4 \left[\frac{x_i (n_{i2} - y_i)}{n_{i.}} \right]}{\sum_{i=1}^4 \left[\frac{y_i (n_{i1} - x_i)}{n_{i.}} \right]}$$

$$\sum_{i=1}^4 \frac{x_i (n_{i2} - y_i)}{n_{i.}} = \frac{4(79-5)}{95} + \frac{2(87-3)}{103} +$$

$$\frac{4(90-10)}{108} + \frac{1(82-3)}{97}$$

$$= 8.5242534$$

$$\sum_{i=1}^4 \frac{y_i (n_{i1} - x_i)}{n_{i.}} = \frac{5(16-4)}{95} + \frac{3(16-2)}{103} +$$

$$\frac{10(18-4)}{108} + \frac{3(15-1)}{97}$$

$$= 2.7686319$$

$$\hat{\psi}_{mh} = \frac{8.5242534}{2.7686319} = 3.0788684$$

$$\therefore \hat{\psi}_{mh} = 3.08$$

$$\ln \hat{\psi}_{gc} = \frac{\sum \hat{W}_i \tilde{\Delta}_i^{(w)}}{\sum \hat{W}_i}, \quad \tilde{\Delta}_i^{(w)} \text{ media ponderada}$$

$$\tilde{\Delta}_i^{(w)} = \ln \left[\frac{x_i - 1/2}{n_{i1} - x_i - 1/2} \right] - \ln \left[\frac{y_i - 1/2}{n_{i2} - y_i - 1/2} \right]$$

$$\tilde{\Delta}_1^{(w)} = \ln \left[\frac{5 - 1/2}{79 - 5 - 1/2} \right] - \ln \left[\frac{4 - 1/2}{16 - 4 - 1/2} \right]$$

$$= 1.6036239$$

$$\tilde{\Delta}_2^{(w)} = \ln \left[\frac{2 - 1/2}{16 - 2 - 1/2} \right] - \ln \left[\frac{3 - 1/2}{87 - 3 - 1/2} \right]$$

$$= 1.3113313$$

$$\tilde{\Delta}_3^{(w)} = \ln \left[\frac{4 - 1/2}{18 - 4 - 1/2} \right] - \ln \left[\frac{10 - 1/2}{90 - 10 - 1/2} \right]$$

$$= .7745385$$

$$\tilde{\Delta}_4^{(w)} = \ln \left[\frac{1 - 1/2}{15 - 1 - 1/2} \right] - \ln \left[\frac{3 - 1/2}{82 - 3 - 1/2} \right]$$

$$= .150971$$

$$\begin{aligned} \Sigma \hat{W}_i \tilde{\Delta}_i^{(w)} &= (2.0101989) (1.6036239) \\ & (1.3046074) (1.3113313) \\ & (2.5072196) (.7745385) \\ & (.967188) (.150971) \end{aligned}$$

$$= 7.022331$$

$$\ln \hat{\psi}_{gc} = \frac{7.022331}{6.789214} = 1.0343364$$

$$\hat{\psi}_{gc} = e^{1.0343364} = 2.8132387$$

$$\therefore \boxed{\hat{\psi}_{gc} = 2.8132}$$

$\hat{\psi}_{am1}$ calculada por la solución iterativa del Método de Cornfield extendido que se encuentra en la Tabla No. 3.

$$z_i = \frac{x_i - \hat{E}(x_i | \psi_i)}{\sqrt{\hat{V}(x_i | \psi_i)}}, \quad \chi_{\alpha} = 0, \quad \epsilon = 10^{-4}$$

$$\psi_0 = 3.08$$

$$\psi_1 = 0.9(3.08) = 2.772$$

$$\text{para } \psi_0 = 3.08 \quad z_0 = ?$$

$$3.08(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$2.08\hat{x}_1^2 - 147\hat{x}_1 + 443.52 = 0$$

$$\hat{x}_1 = \frac{147 - 133.86155}{4.16} = 3.1582821$$

$$3.08(80 - 21\hat{x}_2 + \hat{x}_2^2) = 82\hat{x}_2 + \hat{x}_2^2$$

$$2.08\hat{x}_2^2 - 146.68\hat{x}_2 + 246.4 = 0$$

$$\hat{x}_2 = \frac{146.68 - 139.51693}{4.16} = 1.7218913$$

$$3.08(252 - 32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$2.08\hat{x}_3^2 - 174.56\hat{x}_3 + 776.16 = 0$$

$$\hat{x}_3 = \frac{174.56 - 154.96304}{4.16} = 4.7108088$$

$$3.08(60 - 19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$2.08\hat{x}_4^2 - 136.52\hat{x}_4 + 184.8 = 0$$

$$\hat{x}_4 = \frac{136.52 - 130.76764}{4.16} = 1.38278$$

$$\hat{E}(X_1 | m_1 ; \lambda) = 10.973762$$

$$\hat{V}(X_1 | m_1 ; \lambda) = 1.7260707 \quad \hat{V}(X_1 | m_1 ; \lambda) = 6.0523433$$

$$\hat{V}(X_2 | m_2 ; \lambda) = 1.0332796$$

$$\hat{V}(X_3 | m_3 ; \lambda) = 2.4535734 \quad \therefore Z_0 = \frac{11 - 10.973762}{2.4601511} = .0106651$$

$$\hat{V}(X_4 | m_4 ; \lambda) = .8394196$$

Para $\psi_1 = 2.772$

$Z_1 = ?$

$$2.772(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$2.772(80 - 21\hat{x}_2 + \hat{x}_2^2) = 82\hat{x}_2 + \hat{x}_2^2$$

$$1.772 \hat{x}_1^2 - 139.3 \hat{x}_1 + 399.168 = 0$$

$$1.772 \hat{x}_2^2 - 140.212 \hat{x}_2 + 221.76 = 0$$

$$\hat{x}_1 = \frac{139.3 - 128.74466}{5.544} = 1.9039214$$

$$\hat{x}_2 = \frac{140.212 - 134.49004}{5.544} = 1.0321001$$

$$2.772(252 - 32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$2.772(60 - 19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$1.772 \hat{x}_3^2 - 164.704 \hat{x}_3 + 698.544 = 0$$

$$1.772 \hat{x}_4^2 - 130.668 \hat{x}_4 + 166.32 = 0$$

$$\hat{x}_3 = \frac{164.704 - 148.91651}{5.544} = 2.8476709$$

$$\hat{x}_4 = \frac{130.668 - 126.07637}{5.544} = .8282168$$

$$\hat{E}(X_i | m_i; \lambda) = 6.6119092$$

$$\hat{V}(X_i | m_i; \lambda) = 4.6484913$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.3315511$$

$$\hat{V}(X_2 | m_2; \lambda) = .7693645$$

$$\therefore Z_1 = \frac{11 - 6.6119092}{2.156036} = 2.0352586$$

$$\hat{V}(X_3 | m_3; \lambda) = 1.9248882$$

$$\hat{V}(X_4 | m_4; \lambda) = .6226875$$

De (1) para $i = 0$

$$\ln \psi_2 = 1.1249296 - \frac{.011}{2.035 - .011} (1.0195691 - 1.1249296)$$

$$= 1.1255022$$

$$\therefore \psi_2 = 3.0817641$$

De (2)

$$|.011| \text{ y } |2.0352586|$$

$$\therefore \psi_3 = 3.08 \text{ y } z_3 = .011$$

De (3)

$$\left| \frac{3.08-3.082}{3.082} \right| = .00064892 > 10^{-4}$$

Regresar a (1)

Para $\psi_2 = 3.082$

$z_2 = ?$

$$3.082(144-25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$3.082(80-21\hat{x}_2 + \hat{x}_2^2) = 82\hat{x}_2 + \hat{x}_2^2$$

$$2.082\hat{x}_1^2 - 147.05\hat{x}_1 + 443.808 = 0$$

$$2.082\hat{x}_2^2 - 146.722\hat{x}_2 + 246.56 = 0$$

$$\hat{x}_1 = \frac{147.05 - 133.89425}{4.164} = 3.1594027$$

$$\hat{x}_2 = \frac{146.722 - 139.54925}{4.164} = 1.7225621$$

$$3.082(252-32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$3.082(60-19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$2.082\hat{x}_3^2 - 174.624\hat{x}_3 + 776.664 = 0$$

$$2.082\hat{x}_4^2 - 136.558\hat{x}_4 + 184.92 = 0$$

$$\hat{x}_3 = \frac{174.624 - 155.00156}{4.164} = 4.7124016$$

$$\hat{x}_4 = \frac{136.558 - 130.79783}{4.164} = 1.3833261$$

$$\hat{E}(X_i | m_i; \lambda) = 10.977693$$

$$\hat{V}(X_i | m_i; \lambda) = 6.0531428$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.7262878$$

$$\hat{V}(X_2 | m_2; \lambda) = 1.033451$$

$$\therefore z_2 = \frac{11 - 10.977693}{2.4603136} = .00906693$$

$$\hat{V}(X_3 | m_3; \lambda) = 2.4538414$$

$$\hat{V}(X_4 | m_4; \lambda) = .8395626$$

De (1) para $i = 1$

$$\begin{aligned} \ln \psi_4 &= 1.1255787 - \frac{.009}{.011 - .009} (1.1249296 - 1.255787) \\ &= 1.1284998 \quad \therefore \psi_4 = 3.091016 \end{aligned}$$

De (2)

$$|.009| \text{ y } |.011|$$

$$\therefore \psi_5 = 3.091 \text{ y } z_5 = .009$$

De (3)

$$\left| \frac{3.091 - 3.082}{3.082} \right| = .00292018 > 10^{-4}$$

Regresar a (1)

Para $\psi_4 = 3.091$ $z_4 = ?$

$$3.091(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$2.091\hat{x}_1^2 - 147.275\hat{x}_1 + 445.104 = 0$$

$$\hat{x}_1 = \frac{147.275 - 134.04132}{4.182} = 3.1644378$$

$$3.091(252 - 32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$2.091\hat{x}_3^2 - 174.912\hat{x}_3 + 778.932 = 0$$

$$\hat{x}_3 = \frac{174.912 - 155.17481}{4.182} = 4.7195585$$

$$3.091(80 - 21\hat{x}_2 + \hat{x}_2^2) = 82\hat{x}_2 + \hat{x}_2^2$$

$$2.091\hat{x}_2^2 - 146.911\hat{x}_2 + 247.28 = 0$$

$$\hat{x}_2 = \frac{146.911 - 139.69464}{4.182} = 1.7255767$$

$$3.091(60 - 19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$2.091\hat{x}_4^2 - 136.729\hat{x}_4 + 185.46 = 0$$

$$\hat{x}_4 = \frac{136.729 - 130.93369}{4.182} = 1.385774$$

$$\hat{E}(X_i | m_i; \lambda) = 10.995347$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.7272608$$

$$\hat{V}(X_2 | m_2; \lambda) = 1.0342194$$

$$\hat{V}(X_3 | m_3; \lambda) = 2.4550421$$

$$\hat{V}(X_4 | m_4; \lambda) = .8402019$$

$$\hat{V}(X_i | m_i; \lambda) = 6.0567242$$

$$\therefore z_4 = \frac{11 - 10.995347}{2.4610413} = .00189066$$

De (1) $i = 2$

$$\ln \psi_6 = 1.1284947 - \frac{.002}{.009 - .002} (1.1284947 - 1.1284947)$$

$$= 1.1284947$$

$$\therefore \psi_6 = 3.091$$

De (2)

$$|.002| y |.009|$$

$$\therefore \psi_7 = 3.091 \quad y \quad z_7 = .002$$

De (3)

$$\left| \frac{3.091 - 3.091}{3.091} \right| = 0 < 10^{-4}$$

$$\hat{\psi}_{am1} = 3.091$$

Estimación por Intervalos:

al 95% $\chi^2_{.95} = 3.84$

$$\chi_{.95} = 1.9596 \approx 1.96$$

al 99% $\chi^2_{.99} = 6.63$

$$\chi_{.99} = 2.5749 \approx 2.57$$

Método logit (12.2)

$$\psi_1 = \hat{\psi}_g \exp(-\chi \alpha \hat{W}^{-1/2})$$

$$\psi_2 = \hat{\psi}_g \exp(\chi \alpha \hat{W}^{-1/2})$$

$$\hat{W} = 6.789214$$

$$\hat{W}^{-1/2} = \frac{1}{\sqrt{6.789214}} = .3837869, \hat{\psi}_g = 3.27$$

al 95%

$$\begin{aligned}\psi_1 &= 3.27 \exp (-1.96(.3837869)) \\ &= 1.5412095\end{aligned}$$

$$\begin{aligned}\psi_2 &= 3.27 \exp (1.96(.3837869)) \\ &= 6.9379914\end{aligned}$$

$$\therefore \boxed{1.54 \leq \psi \leq 6.94}$$

al 99%

$$\begin{aligned}\psi_1 &= 3.27 \exp (-2.57(.3837869)) \\ &= 1.2195204\end{aligned}$$

$$\begin{aligned}\psi_2 &= 3.27 \exp (2.57(.3837869)) \\ &= 8.7681191\end{aligned}$$

$$\therefore \boxed{1.22 \leq \psi \leq 8.77}$$

Método logit modificado [21, pp.79-80]

$$\psi_i = \hat{\psi}_{gc} \exp (-\chi_\alpha \hat{w}^{-1/2}), \psi_2 = \hat{\psi}_{gc} \exp(\chi_\alpha \hat{w}^{-1/2})$$

al 95%

$$\begin{aligned}\psi_1 &= 2.8132387 (-1.96(.3837869)) \\ &= 1.3259299\end{aligned}$$

$$\psi_2 = 2.8132387 (1.96(.3837869))$$

$$= 5.9688764$$

$$\therefore \boxed{1.325 \leq \psi \leq 5.967}$$

Hay una pequeña discrepancia con el límite superior reportado en el artículo, que es de 6,3 al 99%

$$\begin{aligned} \psi_1 &= 2.8132387 (-2.57(.3837869)) \\ &= 1.0491749 \end{aligned}$$

$$\begin{aligned} \psi_2 &= 2.8132387 (2.57(.3837869)) \\ &= 7.5433676 \end{aligned}$$

$$\therefore \boxed{1.049 \leq \psi \leq 7.543}$$

Hay una pequeña discrepancia con el límite superior reportado en el artículo $1.02 \leq \psi \leq 8.06$

Birch [5]

$$\psi_1 = \hat{\psi}_b \exp \left\{ -\chi_\alpha \left[\frac{1}{\sqrt{\Sigma}} V(X_i | m_i; \lambda = 0) \right] \right\}$$

$$\psi_2 = \hat{\psi}_b \exp \left\{ \chi_\alpha \left[\frac{1}{\sqrt{\Sigma}} V(X_i | m_i; \lambda = 0) \right] \right\}$$

$$\frac{1}{\sqrt{\Sigma}} V(X_i | m_i; \lambda = 0) = .5001058$$

$$\hat{\psi}_b = 4.2186464$$

al 95%

$$\begin{aligned}\psi_1 &= 4.22 \exp\{-1.96(.5001058)\} \\ &= 1.5834844\end{aligned}$$

$$\begin{aligned}\psi_2 &= 4.22 \exp\{1.96(.5001058)\} \\ &= 11.246337\end{aligned}$$

$$\therefore \boxed{1.583 \leq \psi \leq 11.246}$$

al 99%

$$\begin{aligned}\psi_1 &= 4.22 \exp\{-2.57(.5001058)\} \\ &= 1.1671481\end{aligned}$$

$$\begin{aligned}\psi_2 &= 4.22 \exp\{2.57(.5001058)\} \\ &= 15.258047\end{aligned}$$

$$\therefore \boxed{1.1671 \leq \psi \leq 15.258}$$

Goodman [36]*

al 95%

$$\boxed{0.97 \leq \Psi \leq 7.18}$$

al 99%

$$\boxed{0.71 \leq \psi \leq 15.30}$$

* Para mayor información, explicación del método y cálculo consultar en la bibliografía el No. [36]

Método extendido de Cornfield [31].

Usando la solución iterativa de la tabla 3.

Límite de confianza bajo el 95% para los datos de la
Tabla 2.

$$\chi_{\alpha} = 1.96 \quad , \quad \epsilon = 10^{-4}$$

$$\psi_0 = 1.54$$

$$\psi_1 = (0.9)(1.54) = 1.386$$

para saber el valor correspondiente de las Z_i en las ψ_i tenemos que:

$$\psi_i = \frac{\hat{x}_i (n_{i2} + \hat{x}_i - m_i)}{(n_{i1} - \hat{x}_i) (m_i - \hat{x}_i)}$$

$$\hat{E} (X_i | m_i; \lambda) = \hat{x}_i$$

$$\hat{V} (X_i | m_i; \lambda) = \left[\frac{1}{\hat{x}_i} + \frac{1}{m_i - \hat{x}_i} + \frac{1}{n_{i1} - \hat{x}_i} + \frac{1}{n_{i2} + \hat{x}_i - m_i} \right]^{-1}$$

$$\hat{E} (X. | m_i; \lambda) = \sum_{i=1}^4 \hat{E} (X_i | m_i; \lambda)$$

$$\hat{V} (X. | m_i; \lambda) = \sum_{i=1}^4 \hat{V} (X_i | m_i; \lambda)$$

$$Z_i = \frac{x. - \hat{E}(X. | \psi_i) \pm 1/2}{\sqrt{\hat{V}(X. | \psi_i)}} \quad , \quad \begin{array}{l} + \text{ para el límite superior} \\ - \text{ para el límite inferior} \end{array}$$

Para el límite inferior

Para: $\psi_0 = 1.54$ $Z_0 = ?$

$$1.54 = \frac{\hat{x}_1 (79 - \hat{x}_1 - 9)}{(16 - \hat{x}_1) (9 - \hat{x}_1)}$$

$$1.54(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70 \hat{x}_1 + \hat{x}_1^2$$

$$221.76 - 38.5\hat{x}_1 + 1.54 \hat{x}_1^2 = 70 \hat{x}_1 + \hat{x}_1^2$$

$$.54\hat{x}_1^2 - 108.5\hat{x}_1 + 221.76 = 0$$

$$\hat{x}_1 = \frac{108 - 106.2697}{1.08} = 2.0630926$$

$$1.54 = \frac{\hat{x}_2 (87 + \hat{x}_2 - 5)}{(16 - \hat{x}_2) (5 - \hat{x}_2)}$$

$$1.54 (80 - 21 \hat{x}_2 + \hat{x}_2^2) = 82 \hat{x}_2 + \hat{x}_2^2$$

$$123.2 - 32.34\hat{x}_2 + 1.54 \hat{x}_2^2 = 82\hat{x}_2 + \hat{x}_2^2$$

$$.54\hat{x}_2^2 - 114.34\hat{x}_2 + 123.2 = 0$$

$$\hat{x}_2 = \frac{114.34 - 113.17033}{1.08} = 1.0830278$$

$$1.54 = \frac{\hat{x}_3 (90 + \hat{x}_3 - 14)}{(18 - \hat{x}_3) (14 - \hat{x}_3)}$$

$$1.54(252 - 32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$388.08 - 49.28\hat{x}_3 + 1.54\hat{x}_3^2 = 76\hat{x}_3 + \hat{x}_3^2$$

$$.54\hat{x}_3^2 - 125.28\hat{x}_3 + 388.08 = 0$$

$$\hat{x}_3 = \frac{125.28 - 121.88858}{1.08} = 3.140205$$

$$1.54 = \frac{\hat{x}_4 (82 + \hat{x}_4 - 4)}{(15 - \hat{x}_4) (4 - \hat{x}_4)}$$

$$1.54(60 - 19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$92.4 - 29.26\hat{x}_4 + 1.54\hat{x}_4^2 = 78\hat{x}_4 + \hat{x}_4^2$$

$$.54\hat{x}_4^2 - 107.26\hat{x}_4 + 92.4 = 0$$

$$\hat{x}_4 = \frac{107.26 - 106.32555}{1.08} = .8652271$$

$$\hat{E}(X_i | m_i; \lambda) = 7.1535525$$

$$\begin{aligned} \hat{V}(X_1 | m_1; \lambda) &= \left[\frac{1}{\hat{x}_1} + \frac{1}{9 - \hat{x}_1} + \frac{1}{16 - \hat{x}_1} + \frac{1}{79 + \hat{x}_1 - 9} \right]^{-1} \\ &= 1.4004104 \end{aligned}$$

$$\begin{aligned} \hat{V}(X_2 | m_2; \lambda) &= \left[\frac{1}{\hat{x}_2} + \frac{1}{5 - \hat{x}_2} + \frac{1}{16 - \hat{x}_2} + \frac{1}{87 + \hat{x}_2 - 5} \right]^{-1} \\ &= .7950955 \end{aligned}$$

$$\begin{aligned} \hat{V}(X_3 | m_3; \lambda) &= \left[\frac{1}{\hat{x}_3} + \frac{1}{14 - \hat{x}_3} + \frac{1}{18 - \hat{x}_3} + \frac{1}{90 + \hat{x}_3 - 14} \right]^{-1} \\ &= 2.0388823 \end{aligned}$$

$$\hat{V}(X_4 | m_3; \lambda) = \left[\frac{1}{\hat{x}_4} + \frac{1}{14 - \hat{x}_4} + \frac{1}{15 - \hat{x}_4} + \frac{1}{82 + \hat{x}_4 - 14} \right]^{-1}$$
$$= .6417679$$

$$\hat{V}(X. | m_1; \lambda) = 4.8761561$$

$$\therefore z_0 = \frac{11 - 7.1335525 \cdot 5}{2.208202} = 1.5154626$$

Para: $\psi_1 = 1.386$

$z_1 = ?$

$$1.386(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$.386\hat{x}_1^2 - 104.65\hat{x}_1 + 199.584 = 0$$

$$\hat{x}_1 = \frac{104.65 - 103.16717}{.772} = 1.9207654$$

$$1.386(80 - 21\hat{x}_2 + \hat{x}_2^2) = 82\hat{x}_2 + \hat{x}_2^2$$

$$.386\hat{x}_2^2 - 111.106\hat{x}_2 + 110.88 = 0$$

$$\hat{x}_2 = \frac{111.106 - 110.33288}{.772} = 1.0014503$$

$$1.386(252 - 32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$.386\hat{x}_3^2 - 120.352\hat{x}_3 + 349.272 = 0$$

$$\hat{x}_3 = \frac{120.352 - 118.09034}{.772} = 2.929614$$

$$1.386(60 - 19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$.386\hat{x}_4^2 - 104.334\hat{x}_4 + 83.16 = 0$$

$$\hat{x}_4 = \frac{104.334 - 103.71685}{.772} = .79942$$

$$\hat{E}(X_i | m_i; \lambda) = 6.6512498$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.3390201$$

$$\hat{V}(X_2 | m_2; \lambda) = .753373$$

$$\hat{V}(X_3 | m_3; \lambda) = 1.9581052$$

$$\hat{V}(X_4 | m_4; \lambda) = .6073635$$

$$\hat{V}(X_i | m_i; \lambda) = 4.6578618$$

$$\therefore z_1 = \frac{11 - 6.6512498 \cdot 1.5}{2.158208} = 1.7833083$$

De (1) para $i = 0$

$$\ln \psi_2 = .4317824 - \frac{1.5155 - 1.96}{1.7833 - 1.5155} (.3264219 - .4317824)$$

$$= .2569028$$

$$\therefore \psi_2 = 1.2929195$$

De (2)

$$|1.5155 - 1.96| = .4445$$

$$|1.7833 - 1.96| = .1767$$

$$\therefore \psi_3 = 1.386 \quad y \quad z_3 = 1.7833$$

De (3)

$$\left| \frac{1.386 - 1.2929}{1.2929} \right| = 0.0720086 > 10^{-4}$$

Regresar a (1)

$$\text{Para: } \psi_2 = 1.2929 \quad z_2 = ?$$

$$1.2929(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$.2929\hat{x}_1^2 - 102.3225\hat{x}_1 + 186.1776 = 0$$

$$\hat{x}_1 = \frac{102.3225 - 101.25102}{.5858} = 1.8290947$$

$$1.2929(80 - 21\hat{x}_2 + \hat{x}_2^2) = 82\hat{x}_2 + \hat{x}_2^2$$

$$.2929\hat{x}_2^2 - 109.1509\hat{x}_2 + 103.432 = 0$$

$$\hat{x}_2 = \frac{109.1509 - 108.59437}{.5858} = .9500276$$

$$1.2929(252 - 32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$.2929\hat{x}_3^2 - 117.3728\hat{x}_3 + 325.8108 = 0$$

$$\hat{x}_3 = \frac{117.3728 - 115.73528}{.5858} = 2.7953628$$

$$1.2929(60.19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$.2929\hat{x}_4^2 - 102.5651\hat{x}_4 + 77.574 = 0$$

$$\hat{x}_4 = \frac{102.5651 - 102.12108}{.5858} = .75798$$

$$\hat{E}(X_i | m_i; \lambda) = 6.3324651$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.2975891$$

$$\hat{V}(X_2 | m_2; \lambda) = .7256804$$

$$\hat{V}(X_3 | m_3; \lambda) = 1.9031502$$

$$\hat{V}(X_4 | m_4; \lambda) = .5845705$$

$$\hat{V}(X_i | m_i; \lambda) = 4.5109902$$

$$\therefore z_2 = \frac{11 - 6.3324651 - .5}{2.1239092} = 1.9622001$$

De (1) para $i = 1$

$$\ln \psi_4 = .256877 - \frac{1.9622 - 1.96}{1.7833 - 1.9621} (.3264219 - .2568877)$$

$$= .2577043$$

$$\therefore \psi_4 = 1.2939562$$

$$\text{Para } \psi_4 = 1.2940 \quad z_4 = ?$$

$$1.2949(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$.294\hat{x}_1^2 - 102.35\hat{x}_1 + 186.336 = 0$$

$$\hat{x}_1 = \frac{102.35 - 101.27384}{.588} = 1.8301983$$

$$1.294(80 - 21\hat{x}_2 + \hat{x}_2^2) = 82\hat{x}_2 + \hat{x}_2^2$$

$$.294\hat{x}_2^2 - 109.174\hat{x}_2 + 103.52 = 0$$

$$\hat{x}_2 = \frac{109.174 - 108.61502}{.588} = .9506449$$

$$1.294(252 - 32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$.294\hat{x}_3^2 - 117.408\hat{x}_3 + 326.088 = 0$$

$$\hat{x}_3 = \frac{117.408 - 115.76337}{.588} = 2.7969815$$

$$1.294(60 - 19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$.294\hat{x}_4^2 - 102.586\hat{x}_4 + 77.64 = 0$$

$$\hat{x}_4 = \frac{102.586 - 102.14002}{.588} = .7584772$$

$$\hat{E}(X_1 | m_1; \lambda) = 6.3363019$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.2980993$$

$$\hat{V}(X_2 | m_2; \lambda) = .7260193$$

$$\hat{V}(X_3 | m_3; \lambda) = 1.9038292$$

$$\hat{V}(X_4 | m_4; \lambda) = .5848492$$

$$\hat{V}(X_1 | m_1; \lambda) = 4.512797$$

$$\therefore z_4 = \frac{11 - 6.3363019 - .5}{2.1243345} = 1.9600012$$

De (2)

$$|1.9621 - 1.96| = .0021$$

$$|1.7833 - 1.96| = .1767$$

$$\therefore \psi_5 = 1.2929 \quad y \quad z_5 = 1.9621$$

De (3)

$$\left| \frac{1.2929 - 1.2940}{1.2940} \right| = .000850077 > 10^{-4}$$

Regresar a (1)

De (1) para $i = 2$

$$\begin{aligned} \ln \psi_6 &= .2577382 - \frac{1.96 - 1.96}{1.9621 - 1.96} (.256877 - .2577382) \\ &= .2577382 \end{aligned}$$

$$\therefore \psi_6 = 1.2940$$

De (2)

$$|1.96 - 1.96| = 0$$

$$|1.9621 - 1.96| = .0021$$

$$\therefore \psi_7 = 1.2940$$

De (3)

$$\left| \frac{1.2940 - 1.2940}{1.2940} \right| = 0 < 10^{-4}$$

\therefore Límite inferior al 95% es 1.2940

Para el límite superior

$$\psi_0 = 1.54 \quad z_0 = \frac{11 - 7.1535525 + .5}{2.208202} = 1.9683197$$

$$\psi_1 = (1.1)(1.54) = 1.694 \quad z_1 = ?$$

$$1.69(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$.69\hat{x}_1^2 - 112.25\hat{x}_1 + 243.36 = 0$$

$$\hat{x}_1 = \frac{112.25 - 109.21716}{1.38} = 2.1977073$$

$$1.69(80 - 21\hat{x}_2 + \hat{x}_2^2) = 82x_2 + x_2^2$$

$$.69\hat{x}_2^2 - 117.49 \hat{x}_2 + 135.2 = 0$$

$$\hat{x}_2 = \frac{117.49 - 115.8911}{1.38} = 1.15862$$

$$1.69(252 - 32\hat{x}_3 + \hat{x}_3^2) = 76\hat{x}_3 + \hat{x}_3^2$$

$$.69\hat{x}_3^2 - 130.08 \hat{x}_3 + 425.38 = 0$$

$$\hat{x}_3 = \frac{130.08 - 125.48059}{1.38} = 3.3329083$$

$$1.69(60 - 19\hat{x}_4 + \hat{x}_4^2) = 78\hat{x}_4 + \hat{x}_4^2$$

$$.69\hat{x}_4^2 - 110.11 \hat{x}_4 + 101.4 = 0$$

$$\hat{x}_4 = \frac{110.11 - 108.83174}{1.38} = .9262738$$

$$\hat{E}(X_1 | m_1; \lambda) = 7.6155094$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.4527884$$

$$\hat{V}(X_2 | m_2; \lambda) = .8313773$$

$$\hat{V}(X_3 | m_3; \lambda) = 2.107173$$

$$\hat{V}(X_4 | m_4; \lambda) = .6717464$$

$$\hat{V}(X_1 | m_1; \lambda) = 5.0630851$$

$$Z_1 = \frac{11 - 7.6155094 + .5}{2.25013} = 1.7263405$$

$$\therefore Z_1 = 1.73$$

De (1) para $i = 0$

$$\ln \psi_2 = .4317824 - \frac{1.97 + 1.96}{1.73 - 1.97} (.5270926 - .4317824)$$

$$\ln \psi_2 = 1.9924869 \quad \therefore \psi_2 = 7.3337496$$

De (2)

$$|1.97 + 1.96| = 3.93$$

$$|1.73 + 1.96| = 3.69$$

$$\therefore \psi_3 = 1.694 \quad z_3 = 1.73$$

De (3)

$$\left| \frac{1.694 - 7.334}{7.334} \right| = -0.769021 > 10^{-4}$$

Regresar a (1)

$$\text{Para } \psi_2 = 7.334 \quad z_2 = ?$$

$$7.334(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70\hat{x}_1 + \hat{x}_1^2$$

$$6.334\hat{x}_1^2 - 253.35\hat{x}_1 + 1056.096 = 0$$

$$\hat{x}_1 = \frac{253.35 - 193.46569}{12.668} = 4.7272109$$

$$7.334(80 - 21\hat{x}_2 + \hat{x}_2^2) = 82\hat{x}_2 + \hat{x}_2^2$$

$$6.334\hat{x}_2^2 - 236.014 \hat{x}_2 + 586.72 = 0$$

$$\hat{x}_2 = \frac{236.014 - 202.08283}{12.668} = 2.6784946$$

$$7.334(252 - 32 \hat{x}_3 + \hat{x}_3^2) = 76 \hat{x}_3 + \hat{x}_3^2$$

$$6.334\hat{x}_3^2 - 310.688 \hat{x}_3 + 1848.168 = 0$$

$$\hat{x}_3 = \frac{310.688 - 222.93911}{12.668} = 6.9268144$$

$$7.334(60 - 19\hat{x}_4 + \hat{x}_4^2) = 78 \hat{x}_4 + \hat{x}_4^2$$

$$6.334\hat{x}_4^2 - 217.346 \hat{x}_4 + 440.04 = 0$$

$$\hat{x}_4 = \frac{217.346 - 189.97481}{12.668} = 2.1606556$$

$$\hat{E}(X_i | m_i; \lambda) = 16.493176$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.8259118$$

$$\hat{V}(X_2 | m_2; \lambda) = 1.122366$$

$$\hat{V}(X_3 | m_3; \lambda) = 2.5765719$$

$$\hat{V}(X_4 | m_4; \lambda) = .9116975$$

$$\hat{V}(X_i | m_i; \lambda) = 6.4365472$$

$$z_2 = \frac{11 - 16.493176 + .5}{2.5370351} = - 1.9681147$$

$$\therefore z_2 = - 1.97$$

De (1) para $i = 1$

$$\ln \psi_4 = 1.9925211 - \frac{-1.97 + 1.96}{1.73 + 1.97} (.5270926 - 1.9925211)$$

$$= 1.9885605 \quad \therefore \psi_4 = 7.3050105$$

De (2)

$$|-1.97 + 1.96| = .01$$

$$| 1.73 + 1.96 | = 3.69$$

$$\therefore \psi_5 = 7.334 \quad z_5 = - 1.97$$

De (3)

$$\left| \frac{7.334 - 7.31}{7.31} \right| = .00328317 > 10^{-4}$$

Regresar a (1)

$$\text{Para } \psi_4 = 7.31 \quad z_4 = ?$$

$$7.31(144 - 25\hat{x}_1 + \hat{x}_1^2) = 70 \hat{x}_1 + \hat{x}_1^2$$

$$6.31 \hat{x}_1^2 - 252.75 \hat{x}_1 + 1052.64 = 0$$

$$\hat{x}_1 = \frac{252.75 - 193.16814}{12.62} = 4.7212253$$

$$7.31(80 - 21 \hat{x}_2 + \hat{x}_2^2) = 82 \hat{x}_2 + \hat{x}_2^2$$

$$6.31 \hat{x}_2^2 - 235.51 \hat{x}_2 + 584.8 =$$

$$\hat{x}_2 = \frac{235.51 - 201.75383}{12.62} = 2.6748153$$

$$7.31(252 - 32 \hat{x}_3 + \hat{x}_3^2) = 76 \hat{x}_3 + \hat{x}_3^2$$

$$6.31 \hat{x}_3^2 - 309.92 \hat{x}_3 + 1842.12 = 0$$

$$\hat{x}_3 = \frac{309.92 - 222.61019}{12.62} = 6.9183683$$

$$7.31(60 - 19 \hat{x}_4 + \hat{x}_4^2) = 78 \hat{x}_4 + \hat{x}_4^2$$

$$6.31 \hat{x}_4^2 - 216.89 \hat{x}_4 + 438.6 = 0$$

$$\hat{x}_4 = \frac{216.89 - 189.66024}{12.62} = 2.1576672$$

$$\hat{E}(X_i | m_i; \lambda) = 16.472076$$

$$\hat{V}(X_1 | m_1; \lambda) = 1.8262626$$

$$\hat{V}(X_2 | m_2; \lambda) = 1.1226032$$

$$\hat{V}(X_3 | m_3; \lambda) = 2.5769701$$

$$\hat{V}(X_4 | m_4; \lambda) = .9119124$$

$$\hat{V}(X_i | m_i; \lambda) = 6.4377484$$

$$z_4 = \frac{11 - 16.472076 + .5}{2.5372718} = - 1.959615$$

$$\therefore z_4 = - 1.96$$

De (1) para $i = 2$

$$\ln \psi_6 = 1.9892433 - \frac{-1.96 + 1.96}{-1.97 + 1.96} (1.9925211 - 1.9892433)$$

$$= 1.9892433 \quad \therefore \psi_6 = 7.31$$

De (2)

$$| -1.96 + 1.96 | = 0$$

$$| -1.97 + 1.96 | = 0.01$$

$$\therefore \psi_7 = 7.31 \quad z_7 = -1.96$$

De (3)

$$\left| \frac{7.31 - 7.31}{7.31} \right| = 0 < 10^{-4}$$

\therefore Límite superior al 95% es 7.31

$$\therefore \boxed{1.29 \leq \psi \leq 7.31}$$

Al 99% usando la iteración de la Tabla 3 se obtiene:

$$1.02 \leq \psi \leq 9.25$$

Tabla 3.

Solución iterativa del Método de Cornfield extendido

Sea $Z_i = \frac{x_i - E(X_i | \psi_i) \pm 1/2}{\sqrt{\hat{V}(X_i | \psi_i)}}$, + para el límite superior.
 - para el límite interior.

$$\ln \psi_{2i+2} = \ln \psi_{2i} - \frac{Z_{2i} \pm X_\alpha}{Z_{2i+1} - Z_{2i}} (\ln \psi_{2i+1} - \ln \psi_{2i}), i=0,1,\dots \quad (1)$$

Compare $|Z_{2i} \pm X_\alpha|$ y $|Z_{2i+1} \pm X_\alpha|$, el menor es llamado Z_{2i+3} con su correspondiente ψ_{2i+3} (2)

Si $\left| \frac{\psi_{2i+3} - \psi_{2i+2}}{\psi_{2i+2}} \right| < \epsilon$ (preespecificada) se detiene el pro

ceso y ψ_{2i+2} es el límite; si es $> \epsilon$ regresa a (1) (3)

Para iniciar la iteración

Haga ψ_0 el límite apropiado calculado del método logit

ψ_1 es $(0.9)\psi_0$ para el límite inferior y

ψ_1 es $(1.1)\psi_0$ para el límite superior.

Para la solución $\hat{\psi}_{aml}$, omite $\pm 1/2$ de Z_i , sea $X_\alpha = 0$ en (1) y (2), y haga $\psi_0 = \hat{\psi}_{mh}$ y $\psi_1 = 0.9 \psi_0$.

16. METODOS RELACIONADOS Y APLICACIONES.

Hay varios contextos especiales en los cuales el análisis de una combinación de tablas de 2×2 tienen una relevancia especial.

(i) Comparación Apareada de proporciones. Si $n_{i1} = n_{i2} = 1$, $i = 1, 2, \dots, b$, la distribución de (9.5) se reduce a una distribución binomial con parámetro $e^\lambda / (1 + e^\lambda)$ y el tamaño de muestra es igual a el número de pares, donde $m_i = 1$. Los límites de confianza exactos para λ , ϕ ψ , se encuentran fácilmente de las tablas de la distribución binomial [18]. La prueba de significancia para $\lambda = 0$ está basado en la binomial con $p = 1/2$, y es conocida generalmente como la prueba de McNemar [54]. Nam [55] investigó la sensibilidad de esta prueba cuando es apropiada la prueba equivalente de la sección 8 ajustando por efecto de orden.

La eficiencia relativa de la comparación apareada individualmente de proporciones para una serie independiente fué investigada recientemente por Miettinen [51,53] y Chase [10]. Miettinen [52] también consideró la potencia del caso de "controles múltiples", esto es, cuando $n_{i1} = 1$, $n_{i2} = n_2$, una constante, para $i = 1, 2, \dots, b$.

(ii) Selección de estratos. Aunque en estudios experimentales los estratos son a menudo variables discretas definidas

precisamente tales como sexo y raza en la Tabla 2, puede ocurrir en estudios observacionales que los estratos sean escogidos sobre la base de una covariable continua tal como la edad. Generalmente los estratos son escogidos como intervalos convenientes, tales como los grupos de diez años de edad en el ejemplo de Bliss. Recientemente Cochran [13] desarrolló algunas reglas útiles (cf. Cox [15]) para definir estratos por división de las distribuciones marginales por sus percentiles. Aunque estos procedimientos se basan en una variable dependiente continua, pueden resultar útiles en el análisis de estudios con respuesta cualitativa cuando las muestras son grandes y el método es modificado apropiadamente para varianzas de siguales. Usando simulación en la computadora, Horbach [41b] estudió el efecto de la elección del estrato en el poder empírico de las varias pruebas mencionadas en la sección 14.

(iii) El problema de Bioensayo con respuesta cualitativa. El modelo logístico de la sección 9 puede ser entendido fácilmente para incluir la cuestión de la prueba de potencia relativa de dos drogas usando el análisis logit. Suponiendo que las dosis, d_i , $i = 1, 2, \dots, b$, son idénticas para las dos drogas, simplemente haga

$$\beta_i = \alpha + \beta d_i \quad (16.1)$$

La verosimilitud modificada es entonces una función de un conjunto 3 de estadísticas suficientes conjuntamente: x - y para

λ , $x + y$ para α , y $\Sigma(x_i + y_i)d_i$ para β . Sobre la base de la teoría, las inferencias sobre su diferencia, i.e. λ se deberían basar en la distribución de $X - Y$ condicionada en los valores observados de las otras dos estadísticas. Esta distribución condicional es compleja, pero se puede usar una enumeración para encontrar la prueba exacta óptima [45, p.157 y 21, pp. 69-72]. Los intervalos de confianza para λ pueden encontrarse directamente, y para la potencia relativa, λ/β , usando el enfoque de Cox [20]. Detalles de esto serán reportados subsecuentemente [30]. Como la suposición (16.1) es un caso especial de la suposición general de (9.1), las pruebas e intervalos de confianza derivado sobre esta base son válidos, pero no son óptimos. Mantel [48] sugirió el uso de las pruebas estadísticas de (10.1) para la situación de bioensayo y da un ejemplo numérico.

(iv) Pruebas de aleatorización. Como una alternativa a suponer algún modelo para la proporción de respondentes, Cox [19] también sugirió una prueba de aleatorización sobre los estratos. En su ejemplo, la prueba de aleatorización y la exacta produjeron p 's completamente comparables, 0.02 contra 0.045.

ANEXO I

12

```
10 REM CALCULO DE LAS X(I) USANDO LA ITERACION (6.4)
20 DIM X[100],X2[100]
30 INPUT "CUANTOS VALORES DE X DESEA?",N
31 PRINT
32 PRINT
35 PRINT "   CALCULO DE LAS X(I) USANDO LA ITERACION (6.4)"
36 PRINT
37 PRINT
40 REDIM X[N],X2[N]
50 LET X=4
60 LET N1=16
70 LET N2=79
80 LET M=9
90 LET A=1.9596
95 PRINT "   X(0)=4           LIMITE SUPERIOR           LIMITE INFERIOR"
100 REM CALCULO DE X1 Y X2
110 F=SQR(1/((1/X)+(1/(M-X)))+(1/(N1-X))+(1/(N2+X-M))))
120 X[1]=X+.5+A*F
130 X2[1]=X-.5-A*F
140 PRINT "   X( 1   )   LS=";X[1],"X( 1   )   LI=";X2[1]
150 FOR I=2 TO N
160   F=SQR(1/((1/X[I-1])+(1/(M-X[I-1]))+(1/(N1-X[I-1]))+(1/(N2+X[I-1]&
   J-M))))
165   F2=SQR(1/((1/X2[I-1])+(1/(M-X2[I-1]))+(1/(N1-X2[I-1]))+(1/(N2+&
   X2[I-1]-M))))
170   X[I]=X+.5+A*F
190   X2[I]=X-.5-A*F2
195   PRINT "   X(",I,")   LS=";X[I],"X(",I,")   LI=";X2[I]
200 NEXT I
210 STOP
220 END
```

CUANTOS VALORES DE X DESEA? 15

CALCULO DE LAS X(I) USANDO LA ITERACION (6.4)

X(0)=4		LIMITE SUPERIOR		LIMITE INFERIOR	
X(1)	LS= 7,14993	X(1)	LI= .850069
X(2)	LS= 6,68227	X(2)	LI= 1,83139
X(3)	LS= 6,83971	X(3)	LI= 1,26687
X(4)	LS= 6,79152	X(4)	LI= 1,54095
X(5)	LS= 6,80675	X(5)	LI= 1,39443
X(6)	LS= 6,80198	X(6)	LI= 1,46919
X(7)	LS= 6,80348	X(7)	LI= 1,43008
X(8)	LS= 6,80301	X(8)	LI= 1,45028
X(9)	LS= 6,80316	X(9)	LI= 1,43978
X(10)	LS= 6,80311	X(10)	LI= 1,44522
X(11)	LS= 6,80313	X(11)	LI= 1,44239
X(12)	LS= 6,80312	X(12)	LI= 1,44386
X(13)	LS= 6,80312	X(13)	LI= 1,4431
X(14)	LS= 6,80312	X(14)	LI= 1,44349
X(15)	LS= 6,80312	X(15)	LI= 1,44329

DOLORES

```
10 REM CALCULO DE LAS X(I) USANDO LA ITERACION (6.5)
20 DIM X[100],Y[100],X1[100],Y1[100]
30 INPUT "CUANTOS VALORES DE X DESEA?",N
32 PRINT
33 PRINT
35 PRINT "   CALCULO DE LAS X(I) USANDO LA ITERACION (6.5)"
36 PRINT
37 PRINT
40 REDIM X[N],Y[N],X1[N],Y1[N]
50 LET X=4
60 LET N1=16
70 LET N2=79
80 LET M=9
85 LET A=1.9596
86 PRINT "   X(0)=4   LIMITE SUPERIOR"
90 F0=SQR(1/((1/X)+(1/(M-X))+(1/(N1-X))+(1/(N2+X-M))))
100 X[1]=X+.5+A*F0
110 Y[1]=X[1]
120 X1[1]=X-.5-A*F0
123 Y1[1]=X1[1]
130 PRINT "   X( 1   )   LS=";X[1], "Y( 1   )=";Y[1]
140 REM CALCULO DE "X" Y "Y", LIMITE SUPERIOR
150 FOR I=2 TO N
160   F=SQR(1/(1/(Y[I-1])+(1/(M-Y[I-1]))+(1/(N1-Y[I-1]))+(1/(N2+X[I-1]&
      J-M))))
170   X[I]=X+.5+A*F
180   Y[I]=(I*X[I]+(I-1)*Y[I-1])/(2*I-1)
181   PRINT "   X(";I;")   LS=";X[I], "Y(";I;")=";Y[I]
182 NEXT I
183 PRINT
184 PRINT
185 PRINT
186 PRINT "   X(0)=4   LIMITE INFERIOR"
187 PRINT "   X( 1   )   LI=";X1[1], "Y( 1   )=";Y1[1]
188 REM CALCULO DE "X1" Y "Y1" LIMITE INFERIOR
189 FOR I=2 TO N
200   F1=SQR(1/(1/(Y1[I-1])+(1/(M-Y1[I-1]))+(1/(N1-Y1[I-1]))+(1/(N2+&
      X1[I-1]-M))))
210   X1[I]=X-.5-A*F1
220   Y1[I]=(I*X1[I]+(I-1)*Y1[I-1])/(2*I-1)
230   PRINT "   X(";I;")   LI=";X1[I], "Y(";I;")=";Y1[I]
240 NEXT I
250 STOP
260 END
```

CANTOS VALORES DE X DESEA? 10

CALCULO DE LAS X<I> USANDO LA ITERACION <6.5>

X<0>=4	LIMITE SUPERIOR	
X< 1) LS= 7,14993	Y< 1)= 7,14993
X< 2) LS= 6,68227	Y< 2)= 6,83816
X< 3) LS= 6,79198	Y< 3)= 6,81045
X< 4) LS= 6,80082	Y< 4)= 6,80494
X< 5) LS= 6,80255	Y< 5)= 6,80362
X< 6) LS= 6,80297	Y< 6)= 6,80326
X< 7) LS= 6,80308	Y< 7)= 6,80317
X< 8) LS= 6,80311	Y< 8)= 6,80314
X< 9) LS= 6,80312	Y< 9)= 6,80313
X< 10) LS= 6,80312	Y< 10)= 6,80312

X1<0>=4	LIMITE INFERIOR	
X1< 1) LI= ,850069	Y1< 1)= ,850069
X1< 2) LI= 1,83139	Y1< 2)= 1,50428
X1< 3) LI= 1,41235	Y1< 3)= 1,44912
X1< 4) LI= 1,44038	Y1< 4)= 1,44413
X1< 5) LI= 1,44296	Y1< 5)= 1,44348
X1< 6) LI= 1,4433	Y1< 6)= 1,44338
X1< 7) LI= 1,44335	Y1< 7)= 1,44336
X1< 8) LI= 1,44336	Y1< 8)= 1,44336
X1< 9) LI= 1,44336	Y1< 9)= 1,44336
X1< 10) LI= 1,44336	Y1< 10)= 1,44336

ANEXO II

WORKFILE: POLKRES/PCN105 (06/23/81)

```
10000 INTEGER TABLE(4),I(4)
20000 REAL ALPHA,ALPHAU,PSI1,PSI,PSI2,P,F
30000 F(100)
40000 INTEGER IFCPI,IFAULT
50000 I=0
60000 2
70000 I=I+1
80000 IF(I.EQ.5)GO TO 3
90000 WRITE(7,10)
100000 FORMAT(' CUAL ES EL VALOR',I2,X,' DE LA TABLA')
110000 READ(5,/) TABLE(I)
120000 GO TO 2
130000 3
140000 WRITE(7,20)
150000 20
160000 FORMAT(' CUAL ES EL INTERVALO SUPERIOR')
170000 READ(5,/) ALPHAU
180000 WRITE(7,30)
190000 30
200000 FORMAT(' CUAL ES EL INTERVALO INFERIOR')
210000 READ(5,/) ALPHA
220000 WRITE(7,40)
230000 40
240000 FORMAT(' CUAL ES LA MAXIMA DIMENSION DE LA TABLA')
250000 READ(5,/) IFCPI
260000 CALL LIMITS(TABLE,I,ALPHA,ALPHAU,PSI1,PSI,PSI2,P,F,
270000 - IFCPI,IFAULT)
280000 I=0
290000 4
300000 I=I+1
310000 IF(I.EQ.5)GO TO 5
320000 WRITE(6,50)I,I(1)
330000 40
340000 FORMAT(' EL VALOR NUM.',I2,X,' DE LA TABLA REORIENTADA ES ',I2)
350000 GO TO 4
360000 5
370000 WRITE(6,50)PSI1
380000 50
390000 FORMAT(' EL LIMITE INFERIOR DE CONFIANZA ES',F13.6)
400000 WRITE(6,60)PSI2
410000 60
420000 FORMAT(' EL LIMITE SUPERIOR DE CONFIANZA ES',F13.6)
430000 WRITE(6,70)PSI
440000 70
450000 FORMAT(' COEFICIENTE DE MORIS DE LA TABLA REORIENTADA',F13.6)
460000 WRITE(6,80)P
470000 80
480000 FORMAT(' VALOR DE UNA COLA F DE LA PRUEBA DE FISHER',F13.6)
490000 WRITE(6,90)IFAULT
500000 90
510000 FORMAT(' UBICACION DE FALTA',I13)
520000 STOP
530000 END
540000 SUBROUTINE LIMITS(TABLE,I,ALPHA,ALPHAU,PSI1,PSI,PSI2,P,F,
550000 - IFCPI,IFAULT)
560000 C* ALGOLINDO AS 36 J.D.STATIST.SOC. C. (1971) VOL.20,Nº.1
570000 C*
580000 C* LIMITES EXACTOS DE CONFIANZA PARA DISTRIBUCIONES DE
590000 C* MORIS EN TABLAS DE 2X2 CONECTIVIDAD-CARI, J. J. (1962)
600000 C* LIMITES DE CONFIANZA APROXIMADOS EN " EL RIESGO RELATIVO
610000 C* J.D.STATIST.SOC.5,26,P. 454-463.
620000 C*
630000 INTEGER TABLE,I
640000 COMMON/PSI1/I(2,3,4),H(2,3,4),S,ALPHA2
650000 DIMENSION TABLE(4),I(4),F(100)
660000 DATA F/1/
670000 IFAULT=0
680000 PSI=0.0
690000 P=0.0
700000 DO 1 I=1,4
```



```

59000 1      T(I)=TABLE(I)
60000 C*
61000 C* TABLA ORIENTADA
62000 C*
63000     IP=1
64000     DO 2 I=2,4
65000     IF(T(I).LT.T(IB))IP=I
66000 2     CONTINUE
67000     IC=5-IP
68000     GO TO (4,3,3,4),IP
69000 3     IA=1
70000     IF(T(1).LT.T(4))IA=4
71000     GO TO 6
72000 4     IA=2
73000     IF(T(2).LT.T(3))IA=3
74000 6     IP=5-IA
75000     I1=T(IA)+T(IB)
76000     I2=T(IC)+T(ID)
77000     IF(T(ID)/I2.LE.T(IB)/I1)GO TO 7
78000     T(1)=TABLE(IA)
79000     T(2)=TABLE(IB)
80000     T(3)=TABLE(IC)
81000     T(4)=TABLE(ID)
82000     GO TO 8
83000 7     T(1)=TABLE(IC)
84000     T(2)=TABLE(ID)
85000     T(3)=TABLE(IA)
86000     T(4)=TABLE(IB)
87000 8     PSI1=0.0
88000     PSI2=0.0
89000     T1=T(1)*T(4)
90000     IF(T1.LT.0.0)GO TO 10
91000 C*
92000 C* CALCULO DE COCIENTE DE MOMIOS(RIESGO RELATIVO)
93000 C*
94000     PSI=(T(2)*T(3))/T1
95000 C*
96000 C* CALCULO DE TOTALES MARGINALES
97000 C*
98000     N1=T(1)+T(2)
99000     N2=T(3)+T(4)
100000    N=T(2)+T(4)
101000 C*
102000 C* DETERMINE SI LOS LOG-FACTORIALES SUFICIENTES HAN SIDO
103000 C* EVALUADOS PARA CALCULAR LOG(COEFICIENTES)
104000 C*
105000     L=AMAX1(N1,N2)+1
106000     IF(L.GT.NE(TP))GO TO 13
107000     IF(L.LE.NE)GO TO 10
108000     F(1)=0.0
109000     IL=NE+1
110000     NF=L
111000     DO 9 I=IL,NF
112000     DP=I-1
113000     F(I)=F(I-1)+ALOG10(DP)
114000 9     IF((T(2).EQ.0).OR.(ALPHA1.LE.0.0))GO TO 11
115000     ALPHA2=ALPHA1
116000     I1=T(2)
117000     I2=N
118000     I3=0
119000     I4=I1-1

```

1200
1210
1220
1230
1240
1250
1260
1270
1280
1290
1300
1310
1320
1330
1340
1350
1360
1370
1380
1390
1400
1410
1420
1430
1440
1450
1460
1470
1480
1490
1500
1510
1520
1530
1540
1550
1560
1570
1580
1590
1600
1610
1620
1630
1640
1650
1660
1670
1680
1690
1700
1710
1720
1730
1740
1750
1760
1770
1780
1790
1800

```
C*  
C* PARA EL LIMITE INFERIOR  
C*  
CALL SOL(F,PSI1)  
11 I1=0  
I2=I(2)  
I3=I2+1  
I4=I  
IF(ALPHA2.LE.0.0)GO TO 12  
ALPHA2=ALPHA1  
C*  
C* PARA EL LIMITE SUPERIOR  
C*  
CALL SOL(F,PSI2)  
12 ALPHA2=0.0  
S=1.0  
I=FUN(F,1.0)  
GO TO 100  
13 IF ALL I=L  
GO TO 100  
14 IFAULT=-1  
100 RETURN  
END  
SUBROUTINE SOL(F,H)  
C*  
C* ALGORITMO AS 36.1 J.R.STATIST.SOC.,C(1971) VOL. NO. 1  
C* ENCONTRAR EL TAL QUE F(H)=0, CON H ENTRE 0 Y 1, POR  
C* INTERVALOS DIVIDIDOS EN DOS  
C*  
DIMENSION F(100),Z(3)  
COMMON/PRK1/I1,I2,I3,I4,N1,N2,H,S,ALPHA2  
DATA ACC/1.0E-5/  
S=1.0  
DELTA=0.5  
H=0.5  
1 Z(1)=FUN(F,1.0E-16)  
Z(3)=FUN(F,1.0)  
IF(Z(1)*Z(3).LT.0.0)GO TO 2  
S=-1.0  
GO TO 1  
2 Z(2)=FUN(F,H)  
DELTA=DELTA*0.5  
IF(Z(1)*Z(2).GT.0.0)GO TO 3  
Z(3)=Z(2)  
H=H-DELTA  
GO TO 4  
3 Z(1)=Z(2)  
H=H+DELTA  
4 IF(ABS(Z(1)-Z(3)).GT.ACC) GO TO 2  
IF(S.LT.0.0)H=1.0/H  
RETURN  
END  
FUNCTION FUN(F,FV)  
C*  
C* ALGORITMO 36.2 J.R.STATIST.SOC.,C(1971) VOL.20,NO.1  
C* EVALUACION DE LA FUNCION PARA LIMITES DE CONFIANZA  
C*  
DIMENSION D(2),F(100)  
COMMON/PRK1/I1,I2,I3,I4,N1,N2,H,S,ALPHA2  
DATA DPC/307.0/  
G=S*ALOG10(FV)
```

```

181000  DP=0.0
182000  DP1=M+1
183000  DO 1 J=1, M-1
184000  DX=F (M+1)-F (M-J+2)-F (J)+F (M+1)-F (M-N+J)
185000  - F (M-J+2)+FLCAT (J-1)*G
186000  IF (DP.LT.1X)DP=DX
187000  1 CONTINUE
188000  J1=I1+1
189000  J2=I2+1
190000  DO 3 K=1,2
191000  D(K)=0.0
192000  DO 2 J=J1, J2
193000  DX=F (M+1)-F (M-J+2)-F (J)+F (M+1)-F (M-N+J)
194000  - F (M-J+2)+FLCAT (J-1)*G-DP
195000  IF (DX.LT.-DP)GO TO 2
196000  D(K)=D(K)+10.0**DX
197000  2 CONTINUE
198000  J1=I3+1
199000  J2=I4+1
200000  FIN=D (1)/(D (1)+D (2))-ALPHA2
201000  RETURN
202000  END

```

BIBLIOGRAFIA

- [1] Armitage, P. (1966) The chi-square test for heterogeneity of proportions after adjustment for stratification. *Journal of the Royal Statistical Society, Series B*, 28, 150-163; addendum, 29 (1967), 197.
- [2] Bartlett, M.S. (1935) Contingency table interactions. *Journal of the Royal Statistical Society, Series B*, 2, 248-252.
- [3] Bennett, B.M. (1967) Test of Hypotheses concerning matched samples. *Journal of the Royal Statistical Society, Series B*, 29, 468-474.
- [4] Bennett, B.M. (1968) Note on X^2 tests for matched samples. *Journal of the Royal Statistical Society Series B*, 30, 368-370.
- [5] Birch, M.W. (1964) The detection of partial Association, I: The 2 x 2 case. *Journal of the Royal Statistical Society, Series B*, 26, 313-324.
- [6] Birch, M.W. (1965) The detection of partial Association, II: The general case. *Journal of the Royal Statistical Society, Series B*, 27, 111-124.
- [7] Bishop, Y.M.M. (1969) Full contingency tables, logits, and split contingency tables. *Biometrics*, 25, 383-399.
- [8] Bliss, C.I. (1967) *Statistics in Biology, Vol. 1*, 80-85, New York, McGraw-Hill.

- [9] Buehler, R.J. (1959) Some validity criteria for statistical inferences. *The Annals of Mathematical Statistics*, 30, 845-863.
- [10] Chase, G.R. (1968) On the efficiency of matched pairs in Bernoulli trials. *Biometrika*, 55, 365-369.
- [11] Cochran, W.G. (1950) The comparison of percentages in matched samples. *Biometrika*, 37, 256-266.
- [12] Cochran, W.G. (1954) Some methods of strengthening the common χ^2 tests. *Biometrics*, 10, 417-451.
- [13] Cochran, W.G. (1968) The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*, 24, 295-313.
- [14] Cornfield, J. (1956) A statistical problem arising from retrospective studies. *Proceedings of the Third Berkeley Symposium IV*, (J. Neyman, ed.), 135-148, Berkeley, University of California Press.
- [15] Cox, D.R. (1957) Note on grouping. *Journal of the American Statistical Association*, 52, 543-547.
- [16] Cox, D.R. (1958) The regression analysis of binary sequences. *Journal of the Royal Statistical Society, Series B*, 20, 215-242.
- [17] Cox, D.R. (1958) Some problems connected with statistical inference. *The Annals of Mathematical Statistics*, 29, 357-372.
- [18] Cox, D.R. (1958) Two further applications of a model for binary regression. *Biometrika*, 45, 562-565.

- [19] Cox, D.R. (1966) A simple example of a comparison involving quantal data. *Biometrika*, 53, 215-220.
- [20] Cox, D.R. (1967) Fieller's theorem and a generalization. *Biometrika*, 54, 567-572.
- [21] Cox, D.R. (1970) *The Analysis of Binary Data*. London, Methuen.
- [22] Edwards, A.W.F. (1963) The measure of association in a 2×2 table. *Journal of the Royal Statistical Society, Series A*, 126, 109-114.
- [23] Fisher, R.A. (1935) The logic of inductive inference. *Journal of the Royal Statistical Society, Series A*, 98, 39-54.
- [24] Fisher, R.A. (1956) *Statistical Methods and Scientific Inference*. London, Oliver and Boyd.
- [25] Fisher, R.A. (1962) Confidence limits for a cross-product ratio. *Australian Journal of Statistics*, 4, 41.
- [26] Fraser, D.A.S. (1968) *The Structure of Inference*, New York, John Wiley.
- [27] Gart, J.J. (1962) Approximate Confidence limits for the relative risk. *Journal of the Royal Statistical Society, Series B*, 24, 454-463.
- [28] Gart, J.J. (1962) On the combination of relative risks. *Biometrics*, 18, 601-610.
- [29] Gart, J.J. (1969) An exact test for comparing matched proportions in crossover designs. *Biometrika* 56, 75-80.

- [30] Gart, J.J. (1970) Exact tests in quantal bioassay for the logistic model. (Abstract). *The annals of Mathematical Statistics*, 41, 1391.
- [31] Gart, J.J. (1970) Point and interval estimation of the common odds ratio in the combination of 2×2 tables with fixed marginals. *Biometrika*, 57, 471-475.
- [32] Gart, J.J., D.G. Thomas (1970) Results on confidence limits for the odds ratio. (Submitted).
- [33] Gart, J.J., J.R. Zweifel (1967) On the bias of various estimators of the logit and its variance with application to quantal bioassay. *Biometrika*, 54, 181-187.
- [34] Goodman, L.A. (1964) Simultaneous confidence limits for cross-product ratios in contingency tables. *Journal of the Royal Statistical Society, Series B*, 26, 86-102.
- [35] Goodman, L.A. (1964) Simple methods for analyzing three-factor interaction in contingency tables. *Journal of the American Statistical Association*, 59, 319-352.
- [36] Goodman, L.A. (1969) On partitioning chi-square and detecting partial association in three-way contingency tables. *Journal of the Royal Statistical Society, Series B*, 31, 485-498.
- [37] Goodman, L.A. (1970) The multivariate analysis of qualitative data: interactions among multiple classifications. *Journal of the American Statistical Association*, 65, 226-256.

- [38] Grizzle, J.E., C.F. Starmer, G.G. Koch. Analysis of categorical data by linear models. *Biometrics*, 25, 489-504.
- [39] Haldane, J.B.S. (1940) The mean and variance of the moments of X^2 , when used as a test of homogeneity, when expectations are small. *Biometrika*, 29, 133-143.
- [40] Hannan, J., W.L. Harkness (1963) Normal approximation to the distribution of two independent binomials, conditional on a fixed sum. *The Annals of Mathematical Statistics*, 34, 1593-1595.
- [41] Harkness, W.L. (1965) Properties of the extended hypergeometric distribution. *The Annals of Mathematical Statistics*, 36, 938-945.
- [41a] Horbach, L. (1967) Die Anwendung von Standardisierungsverfahren bei der Auswertung therapeutischer Vergleichsreihen. *Arzneimittel-Forschung (Drug Research)*, 17, 1279-1288.
- [41b] Horbach, L. (1968) Zuteilungsprobleme für therapeutische Vergleiche. *Verhandlungen der Deutschen Gesellschaft für innere Medizin*, 74, 193-196.
- [42] Innes, J.R.M., B.M. Ulland, M.G. Valerio, L. Petrucelli, L. Fishbein, E.R. Hart, A.J. Pallotta, R.R. Bates, H.L. Falk, J.J. Gart, M. Klein, I. Mitchell, J. Peters (1969) Bioassay of pesticides and industrial chemicals for tumorigenicity in mice: a preliminary note. *Journal of the National Cancer Institute*, 42, 1101-1114.
- [43] Jowett, G.H. (1964) The harmonic standardization of

comparisons between success rates in two heterogeneous groups of patients. *Biometrics*, 20, 849-858.

- [44] Kincaid, W.M. (1962) The combination of 2 x m contingency tables. *Biometrics*, 18, 224-228.
- [45] Lehmann, F.L. (1959) *Testing Statistical Hypotheses*, New York, Wiley.
- [46] Lilienfield, A.M.E. Pederson, J.E. Dowd (1967) *Cancer Epidemiology: Methods of Study*. Baltimore. John Hopkins Press, 131-134.
- [47] Lindley, D.V. (1969) Review of "The Structure of Inference" by D.A.S. Fraser, *Biometrika*, 56, 453-456.
- [48] Mantel, N. (1963) Chi-square tests with one degree of freedom: extension of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, 58, 690-700.
- [49] Mantel, N. (1966) Models of complex contingency tables and polychotomous dosage response curves. *Biometrics*, 22, 83-95.
- [50] Mantel, N.W. Haenszel (1959) Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719-748.
- [51] Miettinen, O.S. (1968) The matched pairs design in the case of all-or-none responses. *Biometrics*, 24, 339-352.
- [52] Miettinen, O.S. (1969) Individual matching with multiple

controls in the case of all-or-none responses. *Biometrics*, 25, 339-355.

- [53] Miettinen, O.S. (1970) Estimation of relative risk for individually matched series. *Biometrics*, 26, 75-86.
- [54] Mosteller, F. (1952) Some statistical problems in measuring the subjective response to drugs. *Biometrics*, 8, 220-226.
- [55] Nam, J. (1970) On two tests for comparing matched proportions. (Submitted for publication).
- [56] Norton, H.W. (1945) Calculation of chi-square for complex contingency tables. *Journal of the American Statistical Association*, 40, 251-258.
- [57] Odoroff, C.L. (1970) A comparison of minimum logit chi-square estimation and maximum likelihood estimation in $2 \times 2 \times 2$ and $2 \times 2 \times 3$ contingency tables: tests for interactions. *Journal of the American Statistical Association*, 65, 1617-1631.
- [58] Plackett, R.L. (1969) Multidimensional contingency tables: A survey of models and methods. *Bulletin of the International Statistical Institute*, 40, 131-142, London.
- [59] Radhakrishna, S. (1965) Combination of results from several 2×2 contingency tables. *Biometrics*, 21, 86-98.
- [60] Rasch, G. (1960) *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen. Nielson and Lydiche.

- [61] Roy, S.N., M.A. Kastenbaum (1956) On the hypothesis of no "interaction" in a multiway contingency table. *The Annals of Mathematical Statistics*, 27, 749-757.
- [62] Snedecor, G.W., W.G. Cochran (1967) *Statistical Methods*, 6th ed., Ames. Iowa, Iowa State Press.
- [63] Stevens, W.L. (1951) Mean and variance of an entry in a contingency table. *Biometrika*, 38, 468-470.
- [64] Sverdrup, E. (1966) The present state of the decision theory and the Neyman-Pearson theory. *Review of the International Statistical Institute*, 34, 309-333.
- [65] Thomas, D.G. (1971) Exact limits for the odds ratio in a 2×2 table. *Applied Statistics*, 20 in press.
- [66] Thomas, D.G. (1970) Main effects in $2 \times J \times K$ contingency tables (submitted for publication).
- [67] Thomas, D.G. (1970) Interaction in $2 \times J \times K$ contingency tables (submitted for publication).
- [68] Thomas, D.G. (1970) Exact test for the combination of 2×2 tables. (In preparation).
- [69] Woolf, B. (1955) On estimating the relation between blood group and disease, *Annals of Human Genetics*, 19, 251-253.
- [70] Zelen, M. (1970) Exact significance tests for contingency tables imbedded in a 2^n classification. *Proceedings of the Sixth Berkeley Symposium*, Berkeley University of California Press, in press.

- [71] Zelen, M. (1971) The analysis of several 2×2 contingency tables. *Biometrika*, 58, in press.