110

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

La interpretación del parámetro de asociación del Modelo logarítmico-lineal para dos variables categóricas

T E S I S

QUE PARA OBTENER EL TÍTULO DE

A C T U A R I O

PRESENTA

GUSTAVO RAÚL SCHINCA MUÑOZ

DIRECTOR DE TESIS

M. en C. INOCENCIO RAFAEL MADRID RÍOS

México, D.F. 2002

VNIVERSDAD NACIONAL
AVTONOMA DE
MEXICO

M. EN C. ELENA DE OTEYZA DE OTEYZA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunico a usted que hemos revisado el trabajo escrito:
"LA INTERPRETACIÓN DEL PARÁMETRO DE ASOCIACIÓN DEL MODELO LOGARÍTMICO-LINEAL PARA DOS VARIABLES CATEGÓRICAS"
realizado por GUSTAVO RAÚL SCHINCA MUÑOZ

con número de cuenta 9561869-0 , quién cubrió los créditos de la carrera de: ACTUARÍA

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	M. EN C. INOCENCIO RAFAEL MADRID RÍOS
Propietario	DR. JOAQUÍN CURIEL CAÑEDO
Propietario	ACT. SERGIO HUGO DELGADO ALONSO
Suplente	ACT. FELIPE ZAMORA RAMOS
Suplente	ACT. RICARDO VILLEGAS AZCORRA

Consejo Departamental de MATEMÁTICAS

M. EN C. JOSÉ ANTONIO FLORES DÍAZ

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMATICAS

A Mamá y Papá, porque sé que este logro es también de ustedes

Al pueblo de México, por su esfuerzo por mantener esta Universidad, que más que mi maestra es madre de todos mis actos

ÍNDICE

CAPÍTULO 1

VARIABLES CATEGÓRICAS

1.1 Variables categóricas

Cuando cuantificamos algo, no todos los individuos u objetos producen la misma medida. Por esta razón cuando algo es cuantificado comunmente es llamado variable.[1]

Una variable categórica es aquella cuya escala de medición consiste en un conjunto de categorías. Por ejemplo, la tendencia política puede ser medida como "liberal", "moderada" o "conservadora"; el hábito de fumar puede ser medido usando las categorías "no fuma", "ex-fumador" y "fumador"; y la recuperación de una operación puede ser "completamente recuperado", "mejoría relativa" y "no se ha recuperado".[2]

Hay muchos tipos de variables categóricas. A continuación veremos las clasificaciones que establece Alan Agresti[3] que coinciden casi en su totalidad con las que hace David J. Sheskin[4].

1.1.1 Clasificación Variable Respuesta/Variable Explicativa

La mayoría de los estudios estadísticos hacen la distinción entre variable respuesta (ó "dependiente") y las variables explicativas (ó "dependientes"). Por ejemplo los modelos de regresión describen como la distribución de una respuesta continua, por ejemplo el tiempo de supervivencia tras un transplante de corazón, varía de acuerdo a los niveles de las variables explicativas, por ejemplo la edad ó el nivel de colesterol en la sangre.

En algunos estudios las variables respuesta son categóricas y se utilizan métodos específicos cuando las variables explicativas también son categóricas.

1.1.2 Clasificación según la escala de medición

Las variables cuyas categorías no pueden ser ordenadas son llamadas **nominales**. Algunos ejemplos de variables nominales son la preferencia religiosa (Católico, Judío, Protestante, Otra), el medio de transporte (automóvil, camión, metro, bicicleta, otro), tipo de vivienda (casa sola, apartamento, condominio, otra), raza, género y estado marital. Para las variables nominales el orden en que se en listen las categorías es irrelevante para el análisis estadístico.

Muchas variables tienen categórias que pueden ser ordenadas, dichas variables son llamadas **ordinales**. Algunos ejemplos de variables ordinales son el tamaño de un automóvil (subcompacto, compacto, mediano, grande), clase social (alta, media, baja), opinión sobre el aborto (en contra totalmente, en contra, a favor, totalmente a favor), apreciación de la creatividad de alguna compañía (muy pobre, regular, muy bueno), el diagnóstico de arterioesclerosis en un paciente (con certeza, probablemente, improbable, definitivamente no). Existe claramente un orden entre las categorías de las variables ordinales pero lo que no es muy claro ó desconocido es la distancia entre estas. Mientras podemos concluir que una persona clasificada como "moderado" es más liberal que otra persona clasificada como "conservadora", no podemos cuantificar que tan liberal es esa persona.

Una variable de **intervalo** es aquella en la cual sí puede ser cuantificada la distancia entre cualesquiera dos niveles de sus escala de medición. Por ejemplo, la presión sanguínea, la vida útil de un televisor, la duración de la condena a prisión, ingreso y edad son variables de intervalo.

Existe una jerarquía entre estos tres tipos de variables categóricas. Las variables de intervalo son las de orden superior, las variables ordinales les siguen y, finalmente, las nominales. Los métodos estadísticos diseñados para variables de un tipo puede ser usados también con variables de jerarquía mayor, pero no así con las de menor jerarquía. Por ejemplo, las técnicas estadísticas para variables ordinales pueden ser utilizadas también con variables de intervalo (valiéndose del orden existente entre sus categorías mas pasando por alto las distancias que hay entre estas); pero no pueden ser utilizadas con variables nominales, pues no tiene sentido el orden en las categorías de este tipo de variables. Normalmente, lo mejor es aplicar la técnica estadística apropiada para cada tipo de variable categórica.

La escala de medición que utilicemos para una variable determina la técnica estadística que se utilizará para su estudio. Por ejemplo, la variable "educación" es nominal si la cuantificamos según tipo de educación, tales como "escuela pública" y "escuela privada"; pero es ordinal si la cuantificamos según nivel educativo, tales como "primaria", "secundaria", "preparatoria", "universidad" y "post-grado"; y es una variable de intervalo si la cuantificamos según los años de estudio al utilizar valores naturales 0,1,2,....

1.1.3 Clasificación Variables Continuas/Variables discretas

Las variables están clasificadas en continuas y discretas, según la cantidad de valores que pueden tomar. De hecho, la medición de todas las variables se da en forma discreta debido a las limitaciones de los instrumentos de medición.

La distinción entre variables continuas y discretas se da, en la práctica, debido a que unas pueden tomar muchos valores y otras únicamente pueden tomar pocos valores relativamente. Por ejemplo, la gente que se dedica a la Estadística con frecuencia trata las variables discretas que pueden tomar una cantidad grande de valores como si fuesen continuas, utilizándolas en modelos de regresión y en otros métodos que asumen respuestas continuas.

1.1.4 Clasificación Variable Cuantitativa/Variable Cualitativa

Las variables nominales son cualitativas, pues sus niveles difieren en calidad más no en cantidad, en tanto que las variables de Intervalo son cuantitativas ya que los distintos niveles tienen cantidades diferentes de la característica de interés.

El lugar que ocupan las variables ordinales en la clasificación cuantitativa/cualitativa es confusa. Con frecuencia son tratadas como cualitativas, siendo analizadas con métodos propios para variables nominales. Pero en muchos sentidos, las variables ordinales se asemejan más a las variables de intervalo que a las variables nominales. Estas poseen varias características cuantitativas: cada nivel posee una mayor o menor magnitud de la característica en comparación con los demás niveles; y, aunque no siempre es posible medir, usualmente existe una variable continua en el fondo. La clasificación del prejuicio racial (ninguno, poco, mucho) es una medida burda de una característica continua.

Aunque no está permitido, estrictamente hablando, las personas que se dedican a la estadística con frecuencia aprovechan la naturaleza cuantitativa de las variables ordinales asignándoles calificaciones númericas a las categorías. El propósito de calificar es el de aproximar las distancias relativas entre las categorías. Esto requiere de buen juicio y asesoría por parte de los investigadores que utilizan la escala de medición, esto puede brindar beneficios debido a la variedad de métodos disponibles para el análisis de este tipo de datos.

1.2 Tablas de contingencia

Sean X e Y dos variables respuesta categóricas, X con I categorías e Y con J categorías. Cuando clasificamos los individuos en ambas variables, existen IJ clasificaciones posibles. Las respuestas (x,Y) de un individuo aleatoriamente escogido de una población tienen una distribución de probabilidad. Podemos ordenar dicha distribución en una tabla rectangular con I renglones para las categorías de X y J columnas para las categorías de Y. Las casillas de esta tabla representan los IJ posibles clasificaciones. Sus probabilidades son $\{p_{ij}\}$, donde p_{ij} denota la probabilidad de que (X,Y) este en la casilla del renglón i y de la columna j. Cuando las casillas contienen la frecuencia absoluta, la tabla es llamada Tabla de contingencia, término introducido por Karl Pearson en 1904. Una tabla de contingencia con I renglones y J columnas es llamada tabla I por J ó tabla IxJ.[5]

David J. Sheskin presenta el modelo general para una Tabla de contingencia IxJ[6].

$$
\begin{array}{ccccccccc}
 & C_1 & C_2 & \cdots & C_j & \cdots & C_J & & \\
R_1 & O_{11} & O_{12} & \cdots & O_{1j} & \cdots & O_{1J} & & O_{1+} \\
R_2 & O_{21} & O_{22} & \cdots & O_{2j} & \cdots & O_{2J} & & O_{2+} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\
R_i & O_{i1} & O_{i2} & \cdots & O_{ij} & \cdots & O_{iJ} & & O_{i+} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\
R_I & O_{I1} & O_{I2} & \cdots & O_{Ij} & \cdots & O_{IJ} & & O_{I+} \\
 & & & & & & & & \\
 & O_{+1} & O_{+2} & \cdots & O_{+j} & \cdots & O_{+J} & & n
\end{array}
$$

Hay un total de n observaciones en la tabla.
La notación O_{ij} representa el número de observaciones (es decir, individuos u objetos) existentes en la intersección del i-ésimo renglón y la j-ésima columna. En tanto que O_{i+} y O_{+j} representan el número de observaciones en el i-ésimo y en la j-ésima columna respectivamente.

1.3 La prueba χ^2 (ji-cuadrada) de independencia

La prueba χ^2 de independencia es empleada cuando una única muestra es clasificada en dos dimensiones (ó variables). Se asume que la muestra es seleccionada aleatoriamente de la población que representa. Una de las variables comprende I categorías (con $I \geq 2$), las cuales son representadas por los I renglones de la tabla de contingencia, mientras que la segunda variable comprende J categorías (con $J \geq 2$), las cuales son representadas por las J columnas de la tabla de contingencia. La prueba χ^2 de independencia evalúa la hipótesis de que ambas variables son independientes entre sí, donde independencia significa que no hay manera de predecir, con determinada probabilidad, la categoría (de la primera variable) a que pertenecerá una observación dado que conocemos a que categoría pertenece, según la clasificación establecida por la segunda variable[7].

Otros supuestos necesarios para aplicar la prueba χ^2 de independencia son los siguientes[8]:

a) Las variables categóricas empleadas en el estudio son exclusivas (es decir, no hay intersección) una de la otra .

b) Los datos de la muestra provienen de una selección aleatoria e independiente de n observaciones. Esto último implica que cada individuo u objeto sólo puede estar representada una vez en la muestra.

c) La frecuencia esperada en cada celda de la tabla de contingencia es mayor o igual a 5.

d) El número de observaciones por categoría, ya sea en los renglones ó en las columnas, no es determinado por el investigador con anterioridad a la recolección de los datos.

Veamos a continuación un ejemplo de la prueba χ^2 de independencia que presenta David J. Sheskin[9]:

Un investigador desea determinar si existe relación entre la personalidad (introvertido/extrovertido) y la filiación política. Doscientas personas fueron reclutadas para participar en el estudio. A cada individuo se le aplicó una prueba de personalidad para poder clasificarlo, ya fuera como introvertido o como extrovertido. Además, a cada individuo se le pidió que indicase si era Demócrata ó Republicano. En la siguiente Tabla de contingencia 2×2 se presentan los resultados de la prueba.

	Demócrata	*Republicano*	*Suma*
Introvertido	30	70	100
Extrovertido	60	40	100
Suma	90	110	200

¿Implican estos datos que hay una relación entre filiación política y la personalidad? La prueba indicada para saberlo es la χ^2 de independencia pues el estudio cumple con los supuestos necesarios, a saber:

a) El estudio realiza sólamente una muestra.

b) Las categorías de ambas variables son exclusivas una de la otra.

c) El investigador no fijó la cantidad de individuos por categoría en ninguna de las dos variables con antelación a la recolección de la información.

La hipótesis que haremos es que ambas variables (tipo de personalidad y filiación política) son independientes, la cual que equivale a la igualdad indicada por H_o, que llamaremos hipótesis nula. En tanto que, si rechazamos la hipótesis de indepencia (H_o), la hipótesis que aceptamos es H_1, que llamaremos hipótesis alternativa.

$$H_o\text{: } p_{ij} = p_{i+} \cdot p_{+j} \quad \text{vs.} \quad H_1\text{: } p_{ij} \neq p_{i+} \cdot p_{+j}$$

La distribución de probabilidad $\{p_{ij}\}$ es la distribución conjunta de X e Y. En tanto que las distribuciones marginales son los totales obtenidos al sumar por renglón y por columna las probabilidades conjuntas, estas son denotadas con $\{p_{i+}\}$ y $\{p_{+j}\}$ para la variable renglón y la variable columna respectivamente, donde el sub-índice "+" denota la suma sobre el sub-índice que reemplaza, esto es:

$$p_{i+} = \sum_{j=1}^{J} p_{ij} \qquad \text{y} \qquad p_{+j} = \sum_{i=1}^{I} p_{ij},$$

Las cuales satisfacen $\sum_{i=1}^{I} p_{i+} = \sum_{j=1}^{J} p_{+j} = \sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij} = 1$.

Para realizar la prueba χ^2 de independencia, tenemos que comparar el valor observado en cada casilla O_{ij} con la frecuencia esperada $E_{ij} = E(O_{ij}) = \frac{O_{i+} \cdot O_{+j}}{n}$ que corresponde al supuesto de independencia (H_o: $p_{ij} = p_{i+} \cdot p_{+j}$). Dicha comparación se realiza mediante

$$\chi^2 = \sum_{i=1}^{I}\sum_{j=1}^{J} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}.$$

La siguiente tabla resume los cálculos necesarios para calcular χ^2.

Casilla	O_{ij}	E_{ij}	$(O_{ij} - E_{ij})$	$(O_{ij} - E_{ij})^2$	$\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$
I/D	30	45	-15	225	5.00
I/R	70	55	15	225	4.09
E/D	60	45	15	225	5.00
E/R	40	55	-15	225	4.09

$$\chi^2 = \sum_{i=1}^{I}\sum_{j=1}^{J} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 5.00 + 4.09 + 5.00 + 4.09 = 18.18$$

El valor χ^2 obtenido es evaluado con la tabla de la distribución ji-cuadrada. Los grados de libertad para el análisis de independencia es gl = (I-1)(J-1). Ya que en nuestro caso I = 2 y J = 2, tenemos que df = 1. Además, el percentil 95 de la distribución ji-cuadrada con df=1 es $\chi^2_{.95} = 3.84$. Dado que $\chi^2 = 18.18$ es mayor que dicho percentil, podemos rechazar la hipótesis nula. Esto permite al investigador concluir que el tipo de personalidad está asociado con la filiación política. Es importante notar que esto no equivale a una relación causa-efecto, tan sólo se puede concluir que existe una correlación (no nula) entre ambas variables.

1.4 El cociente de ventajas

El cociente de ventajas (atribuido a Cornfield (1951)) es una medida de asociacion que, aunque puede ser aplicada en tablas de contingencia de cualquier dimensión, es más fácil de interpretar en las tablas de contingencia 2×2[10].

1.4.1 Tablas de contingencia 2×2.

Antes de considerar el concepto de cociente de ventajas, será de utilidad explicar el concepto de ventaja.

El cociente

$$p_{1|1} = \frac{p_{11}}{p_{1+}}$$

es llamado probabilidad condicional, representa la probabilidad que tiene una observación en el renglón 1 de pertenecer también a la columna 1 y el cociente

$$p_{2|1} = \frac{p_{12}}{p_{1+}}$$

representa la probabilidad que tiene una observación en el renglón 1 de pertenecer también a la columna 2.

Así, al realizar el cociente

$$\Omega_1 = \frac{p_{1|1}}{p_{2|1}}$$

estamos calculando la ventaja que tiene, en el primer renglón, la columna 1 sobre la columna 2.

De manera análoga para el segundo renglón, la ventaja que tiene la columna 1 sobre la columna 2 está representada por el cociente

$$\Omega_2 = \frac{p_{1|2}}{p_{2|2}}$$

donde $p_{1|2} = \dfrac{p_{21}}{p_{2+}}$ y $p_{2|2} = \dfrac{p_{22}}{p_{2+}}$.

El cociente

$$\theta = \frac{\Omega_1}{\Omega_2} = \frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}}$$

es llamado el cociente de ventajas.

El cociente de ventajas puede tomar cualquier valor no negativo (es decir, $0 \leq \theta < \infty$) y la independencia de X e Y es equivalente a $\theta = 1$. Cuando $1 < \theta < \infty$ indica que los individuos en el renglón 1 tienen mayor probabilidad que los individuos en el renglón 2 de pertenecer a la columna 1, es decir, $p_{1|1} > p_{1|2}$.

Por ejemplo, si $\theta = 4$, la ventaja de la columna 1 es cuatro veces mayor en el primer renglón que la ventaja de la columna 1 en el segundo renglón (esto no significa que la probabilidad $p_{1|1}$ es cuatro veces mayor que $p_{1|2}$). Cuando $0 < \theta < 1$, la primera columna es menos probable en el renglón 1 que en el renglón 2, es decir, $p_{1|1} < p_{1|2}$. Cuando alguna casilla tiene probabilidad cero, θ es igual a 0 ó a ∞.

Los valores de θ más alejados de 1 , en cualquiera de ambas direcciones, representan niveles de asociación más pronunciados. Es más conveniente, en algunas ocasiones, utilizar $\ln(\theta)$. Así, la independencia entre ambas variables corresponde a $\ln(\theta) = 0$ y cuando tenemos dos valores del logaritmo natural del cociente de ventajas que varían únicamente por el signo, estos representan el mismo nivel de asociación pero en sentido contrario. Por ejemplo $\ln(4) = 1.39$ y $\ln\left(\frac{1}{4}\right) = -1.39$[11].

Ronald R. Christensen[12] menciona que la variable aleatoria

$$z = \frac{\ln\left(\widetilde{\theta}\right) - 0}{SE} \sim N(0,1)$$

puede ser empleada para la prueba de hipótesis

$$\mathrm{H}_o\text{: } \theta = 1 \quad \text{vs.} \quad \mathrm{H}_1\text{: } \theta \neq 1,$$

donde

$$\widetilde{\theta} = \frac{O_{11} \cdot O_{22}}{O_{21} \cdot O_{12}}$$

es el cociente de ventajas muestral

y

$$\mathrm{SE} = \sqrt{\frac{1}{O_{11}} + \frac{1}{O_{12}} + \frac{1}{O_{21}} + \frac{1}{O_{22}}}$$

representa el error estándar.

Para realizar dicha prueba se construye el intervalo de confianza al nivel de significancia $1-\frac{\alpha}{2}$:

$$\left(\exp\left[\ln\left(\widetilde{\theta}\right)-SE\cdot Z_{1-\frac{\alpha}{2}}\right],\exp\left[\ln\left(\widetilde{\theta}\right)+SE\cdot Z_{1-\frac{\alpha}{2}}\right]\right),$$

donde $Z_{1-\frac{\alpha}{2}}$ es el percentil $1-\frac{\alpha}{2}$ de la distribución $N(0,1)$.

Para ejemplificar esto, valgámonos del estudio de asociación entre tipo de personalidad (Introvertido/Extrovertido) y filiación política (Demócrata/Republicano) que vimos en la sección anterior.

Recordemos que aceptamos la hipotesis alternativa

$$H_1 : p_{ij} \neq p_{i+}\cdot p_{+j},$$

es decir que existe dependencia (ó asociación) entre el tipo de personalidad y la filiación política.

Pero ¿en qué forma se da dicha asociación? Es una pregunta que con la prueba χ^2 de independiencia no pudimos contestar, pero, en cambio, con el cociente de ventajas, sí podremos hacerlo.

Calculemos, con el nivel de significancia $\alpha = .10$, el intervalo de confianza que recién vimos.

a) $\ln\left(\widetilde{\theta}\right)=\ln\left(\frac{O_{11}\cdot O_{22}}{O_{21}\cdot O_{12}}\right)=\ln\left(\frac{30\cdot 40}{60\cdot 70}\right)=-1.2528$

b) SE $=\sqrt{\frac{1}{O_{11}}+\frac{1}{O_{12}}+\frac{1}{O_{21}}+\frac{1}{O_{22}}}=0.2988$

c) El percentil $Z_{1-\frac{\alpha}{2}}=Z_{0.975}$ de la $N(0,1)$ es 1.96

d) Ahora podemos obtener el intervalo

$$\left(\exp\left[\ln\left(\widetilde{\theta}\right)-SE\cdot Z_{1-\frac{\alpha}{2}}\right],\exp\left[\ln\left(\widetilde{\theta}\right)+SE\cdot Z_{1-\frac{\alpha}{2}}\right]\right)=$$

$$(\exp[-1.2528-0.2928\cdot 1.96],\exp[-1.2528+0.2928\cdot 1.96])=$$

$$(0.1609, 0.5072)$$

Finalmente, podemos concluir que $\theta < 1$ con un 95% de confianza, pues $\theta \in (0.1609, 0.5072)$ con un 95% de confianza.

Lo anterior significa que podemos afirmar, con un 95% de confianza, que las personas con personalidad introvertida tienen mayor probabilidad de ser demócratas que las personas con personalidad extrovertida.

De hecho podemos plantear la conclusión de diversas maneras como, por ejemplo, las personas con personalidad extrovertida tienen mayor probabilidad de ser republicanos que las personas con personalidad introvertida, ó, las personas con personalidad extrovertida tienen menor probabilidad de ser demócratas que las personas con personalidad introvertida, etc.

1.4.2 El cociente de ventajas en tablas de contingencia $I \times J$.

El cociente de ventajas también es de utilidad para analizar tablas de contingencia más grandes que las tablas 2×2. Los cocientes de ventajas para tablas $I \times J$ se calculan a partir de $\binom{I}{2} = \frac{I \cdot (I-1)}{2}$ posibles pares de renglones de la tabla de contingencia en combinación $\binom{J}{2} = \frac{J \cdot (J-1)}{2}$ posibles pares de columnas, los cuales se presentan a continuación[13].

$$\theta_{ij} = \frac{p_{ij} \cdot p_{i+1,\ j+1}}{p_{i,\ j+1} \cdot p_{i+1,\ j}} \quad , \quad i = 1, \ldots, I-1 \ , \ j = 1, \ldots, J-1$$

Aunque el cociente de ventajas, tal como lo mencionamos, puede ser extendido más alla de las tablas 2×2 se vuelve más díficil de interpretar. De cualquier manera, en los casos en que hay más de dos renglones pero hay tan sólo dos columnas, la interpretación sigue siendo relativamente directa[14]. Para ilustrar esto, veamos el siguiente ejemplo.

Un investigador lleva a cabo un estudio con la finalidad de evaluar el efecto del ruido en el comportamiento altruísta. En el experimento participan 300 individuos, los cuales son asignados aleatoriamente a tres condiciones experimentales distintas. A dichos individuos se les aplica una prueba con duración de una hora, a 100 individuos se les expone a un ruido muy escandaloso, el cual, se les dice, proviene de un generador en mal estado. En tanto, otros 100 individuos son expuestos a un ruido moderado y los 100 restantes no son expuestos a ninguna clase de ruido durante la prueba. Al terminar esta parte de la prueba, cada individuo deja el cuarto en que se encontraba y es confrontado con un hombre de mediana edad cuyo brazo está en un cabestrillo, el cual les pide su ayuda para llevar a su carro un pesado paquete. Los resultados de dicha prueba se presentan en la siguiente tabla de contingencia.

	Ayudó	*No ayudó*	*Suma*
Mucho ruido	30	70	100
Ruido moderado	50	50	100
Sin ruido	80	20	100
Suma	160	140	300

La ventajas muestrales por renglón son, respectivamente, $\tilde{\Omega}_1 = \frac{30}{70} = .43$, $\tilde{\Omega}_2 = \frac{50}{50} = 1$ y $\tilde{\Omega}_3 = \frac{80}{20} = 4$.

Entonces la ventaja del resultado "Ayudó" en alguien bajo la condición "sin ruido" es $\frac{\tilde{\Omega}_3}{\tilde{\Omega}_1} = 9.3$ veces mayor que alguien bajo la condición "mucho ruido" y $\frac{\tilde{\Omega}_3}{\tilde{\Omega}_2} = 4$ veces mayor que alguien bajo la condición "ruido moderado". En tanto que la ventaja del resultado "Ayudó" en alguien bajo la condición "ruido moderado" es $\frac{\tilde{\Omega}_2}{\tilde{\Omega}_1} = 2.33$ veces mayor que alguien en bajo la condición "mucho ruido".

Entonces, entre mayor sea el ruido al que es expuesto cada individuo menor es la ventaja muestral de la columna "Ayudó".

1.5 Colapsabilidad

Decimos que una tabla de contingencia es colapsable cuando podemos combinar una o más categorías de alguna de las variables y la relación entre estas que presenta la tabla colapsada no cambia.

Cuando las tablas no son colapsables, el manejar únicamente tablas bi-dimensionales provenientes de sumas de categorías puede llevarnos a conclusiones falsas en cuanto los patrones de interacción entre las variables, tal como vemos en estas tres tablas[15]:

TABLA A

4	2	6
6	3	9

TABLA B

4	8
6	12

TABLA C

4	1	7
6	6	6

La tabla A, sí es colapsable pues las ventajas $\Omega_1 = \frac{4}{6}$, $\Omega_2 = \frac{2}{3}$ y $\Omega_3 = \frac{6}{9}$ indican independencia ya que son constantes, así que podemos sumar las columnas segunda y tercera y la independencia no se verá afectada como vemos en la tabla B, pues $\theta = \frac{4 \times 12}{6 \times 8} = 1$.

En cambio la tabla C no es colapsable ya que $\Omega_1 = \frac{4}{6}$, $\Omega_2 = \frac{1}{6}$ y $\Omega = \frac{7}{6}$ contradicen la independencia y cometeríamos un error si, en vez de estudiarla directamente, sumáramos sus columnas segunda y tercera y analizáramos la tabla B, pues, como vimos, esta sí presenta independencia.

En general podemos identificar las tablas bi-dimensionales que son colapsables cómo las que presentan independencia.

Identificar las tablas colapsables en más de dos dimensiones es una herramienta muy útil para manejar tablas muy grandes, pues podemos compactarlas sin distorsionar la estructura de estas. A continuación veremos un ejemplo que ilustra cómo agregar categorías en tablas tri-dimensionales nos puede llevar a conclusiones erróneas.

La siguiente tabla de contingencia de tres dimensiones, obtenida por Radelet en 1981[16], presenta los resultados que obtuvo en un estudio de los efectos de la raza de un individuo acusado de homicidio en la condena a pena de muerte. Las variables son "pena de muerte", con categorías "sí" y "no", "raza del acusado" y "raza de la víctima(del homicidio)" las cuales tienen las categorías "blanca" y "negra". Los datos provienen de 20 condados de Florida en los que 326 individuos fueron juzgados por homicidio entre 1976 y 1977.

Raza del acusado	*Raza de la víctima*	*Pena de muerte Sí*	*Pena de muerte No*	*Porcentaje pena de muerte*
Blanca	*Blanca*	19	132	12.6 %
	Negra	0	9	0.0 %
Negra	*Blanca*	11	52	17.5 %
	Negra	6	97	5.8 %

Si ignoramos la raza de la víctima, obtendríamos la siguiente tabla de dos dimensiones, la cual llamaremos marginal:

Raza del acusado	*Pena de muerte Sí*	*Pena de muerte No*	*Total*
Blanca	19	141	160
Negra	17	149	166

El porcentaje de acusados de raza blanca que reciben la pena de muerte es $100\left(\frac{19}{160}\right) = 11.87\,\%$, que es mayor que el porcentaje de acusados de raza negra que reciben la pena de muerte, a saber $100\left(\frac{17}{166}\right) = 10.24\,\%$.

En cambio, si consideramos la raza de la víctima tenemos dos tablas de contingencia de dos dimensiones, las cuales llamaremos tablas parciales.

Si la víctima es de raza blanca tenemos la siguiente tabla parcial :

Raza del acusado	*Pena de muerte Sí*	*No*	*Total*
Blanca	19	132	151
Negra	11	52	63

El porcentaje de acusados de raza negra que reciben la pena de muerte es $100\left(\frac{11}{63}\right) = 17.46\,\%$, que es mayor que el porcentaje de acusados de raza blanca que reciben la pena de muerte, a saber $100\left(\frac{19}{151}\right) = 12.58\,\%$.

Si la víctima es de raza negra tenemos la siguiente tabla parcial :

Raza del acusado	*Pena de muerte Sí*	*No*	*Total*
Blanca	0	9	9
Negra	6	97	103

El porcentaje de acusados de raza negra que reciben la pena de muerte es $100\left(\frac{6}{97}\right) = 5.83\,\%$, que es mayor que el porcentaje de acusados de raza blanca que reciben la pena de muerte, a saber $100\left(\frac{0}{9}\right) = 0.00\,\%$.

Si en el estudio consideramos la raza de la víctima, podemos ver que la pena de muerte es en ambos casos más frecuente para los individuos de raza negra.

En consecuencia esta tabla no es colapsable pues la asociación marginal es contraria a las asociaciones parciales, esta situación es llamada Paradoja de Simpson.

En el capítulo siguiente veremos cuáles son las condiciones de colapsabilidad.

REFERENCIAS

[1]Sheskin, David J., Handbook of parametric and non parametric statistical procedures, pág. 3

[2]Agresti, Alan, Categorical Data Analysis, pág. 2

[3]Ibid., págs. 2-4

[4]Sheskin, David J., op.cit., págs. 2-3

[5]Agresti, Alan, op.cit., págs. 8-9

[6]Sheskin, David J., op.cit., pág. 357

[7]Ibid., pág. 358

[8]Ibid., pág. 358

[9]Ibid., pág. 362

[10]Ibid., pág. 402

[11]Agresti, Alan, op.cit., págs. 15-16

[12]Christensen, Ronald R., Log-linear models, citado por Sheskin, David J., op.cit., págs. 406-407

[13]Agresti, Alan, op.cit., pág. 18

[14]Sheskin, David J., op.cit., pág. 405

[15]Bishop, Yvonne M.M. et. al., Discrete Multivariate Analysis, pág. 28

[16]Agresti, Alan, op.cit., págs. 135-136

CAPÍTULO 2

LOS MODELOS LOGARÍTMICO-LINEALES

2.0 Los modelos logarítmico-lineales

Los modelos logarítmico-lineales describen los patrones de asociación entre variables categóricas. Mediante el uso de estos podemos modelar las observaciones en una tabla de contingencia en términos de la asociación entre las variables que la conforman.

En el capítulo anterior vimos que las herramientas para analizar una tabla de contingencia de dos dimensiones son la prueba χ^2 de independencia y el cociente de ventajas. Ahora veremos que dichas herramientas pueden ser expresadas en términos de modelos logarítmico-lineales y que en el caso 2×2 los modelos logarítmico-lineales no nos darán más información, pero para tres o más dimensiones nos permitirán ir más allá del simple modelo de independencia y explorar otros tipos de relaciones entre las variables.

2.1 Modelos para tablas de dos dimensiones

Una manera sencilla de construir un modelo lineal del logaritmo natural de las probabilidades de las celdas de una tabla de contingencia, es por analogía con los modelos de análisis de varianza (ANOVA)[1].

Si definimos $I_{ij} = \ln(p_{ij})$ (bajo el supuesto $p_{ij} > 0$ *para* $i = 1,\ldots,I$ *y para* $j = 1,\ldots J$) entonces podemos hacer la analogía con los modelos ANOVA de la siguiente manera*:

$$i)\quad U = \frac{I_{++}}{IJ} \quad donde \ I_{++} = \sum_{i=1}^{I}\sum_{j=1}^{J} I_{ij}$$

$$ii)\quad U_1(i) = \frac{I_{i+}}{J} - \frac{I_{++}}{IJ} \quad donde \ I_{i+} = \sum_{j=1}^{J} I_{ij}$$

$$iii)\quad U_2(j) = \frac{I_{+j}}{I} - \frac{I_{++}}{IJ} \quad donde \ I_{+j} = \sum_{i=1}^{I} I_{ij}$$

$$iv)\quad U_{12}(i,j) = I_{ij} - (U + U_1(i) + U_2(j))$$

*Esta notación fue establecida por Birch en 1963

Dado que los parámetros $U_1(i)$ y $U_2(j)$ representan las desviaciones del promedio general que es U y los parámetros $U_{12}(i,j)$ representan desviaciones del modelo de independencia $I_{ij} = U + U_1(i) + U_2(j)$ tenemos que

$$\sum_{i=1}^{I} U_1(i) = \sum_{j=1}^{J} U_2(j) = \sum_{i=1}^{I} U_{12}(i,j) = \sum_{j=1}^{J} U_{12}(i,j) = 0$$

Estas restricciones reducen el número de parámetros independientes. El valor $U_1(i)$ difiere para cada una de las I categorías, pero la anterior restricción reducen el número de parámetros independientes a I-1. Análogamente, U_{12} es un arreglo de $I \times J$ valores que suman cero en cada renglón y en cada columna quedando así tan sólo (I-1)(J-1) valores independientes.

Por lo anterior, el modelo logarítmico-lineal

$$I_{ij} = U + U_1(i) + U_2(j) + U_{12}(i,j) \quad \textit{para} \quad i = 1, \ldots, I \ \textit{y para} \ j = 1, \ldots J$$

es llamado "saturado", ya que tiene tantos parámetros independientes como celdas tiene la tabla de contingencia.

Podemos expresar la independencia con el siguiente modelo:

$$I_{ij} = U + U_1 + U_2$$

pues, como veremos más adelante, es equivalente a la siguiente aseveración:

$$\Pr(A_i \cap B_j) = \Pr(A_i)\Pr(B_j) \ \textit{para} \ i = 1, \ldots, I \ \textit{y para} \ j = 1, \ldots J$$

donde A_i representa la i-ésima categoría de la primera variable y B_j la j-ésima categoría de la segunda variable.

Podemos escribir la igualdad anterior de la siguiente manera:

$$p_{ij} = p_{i+} \cdot p_{+j} \ \textit{para} \ i = 1, \ldots, I \ \textit{y para} \ j = 1, \ldots J$$

donde $p_{ij} = \Pr(A_i \cap B_j)$, $p_{i+} = \sum_{j=1}^{J} p_{ij} = \Pr(A_i)$ y $p_{+j} = \sum_{i=1}^{I} p_{ij} = \Pr(B_j)$.

A continuación probaremos estas igualdades.

$$p_{i+} = \sum_{j=1}^{J} p_{ij} = \sum_{j=1}^{J} \Pr(A_i \cap B_j) = \Pr((A_i \cap B_1) \cup (A_i \cap B_2) \cup \cdots \cup (A_i \cap B_J))$$
$$= \Pr\left(A_i \cap \left(\cup_{j=1}^{J} B_j\right)\right) = \Pr(A_i \cap \Omega) = \Pr(A_i)$$

donde Ω representa la población bajo estudio. Cabe mencionar que la tercera igualdad se debe a que las categorías de cualquier variable categórica deben ser excluyentes, es decir, $B_s \cap B_j = \emptyset$ *si* $s \neq j$. Lo anterior lleva a que $(A_i \cap B_s) \cap (A_i \cap B_j) = \emptyset$ *si* $s \neq j$ y por lo tanto a que
$\Pr((A_i \cap B_s) \cup (A_i \cap B_j)) = \Pr(A_i \cap B_s) + \Pr(A_i \cap B_j)$ *si* $s \neq j$. Además, la quinta igualdad se debe a que las variables categóricas son exhaustivas, es decir, $\cup_{j=1}^{J} B_j = \Omega$.

Análogamente obtenemos que

$$p_{+j} = \sum_{i=1}^{I} p_{ij} = \sum_{i=1}^{I} \Pr(B_j \cap A_i) = \Pr((B_j \cap A_1) \cup (B_j \cap A_2) \cup \cdots \cup (B_j \cap A_I))$$
$$= \Pr\left(B_j \cap \left(\cup_{i=1}^{I} A_i\right)\right) = \Pr(B_j \cap \Omega) = \Pr(B_j).$$

q.e.d.

Ahora probaremos que $I_{ij} = U + U_1(i) + U_2(j) \Leftrightarrow p_{ij} = p_{i+} \cdot p_{+j}$.

Primero probaremos que $I_{ij} = U + U_1(i) + U_2(j) \Rightarrow p_{ij} = p_{i+} \cdot p_{+j}$.

Por un lado tenemos que

$$I_{ij} = U + U_1(i) + U_2(j) \Rightarrow p_{ij} = \exp(U + U_1(i) + U_2(j))$$
$$\Rightarrow p_{i+} = \sum_{j=1}^{J} \exp(U + U_1(i) + U_2(j)) = \sum_{j=1}^{J} \exp(U)\exp(U_1(i))\exp(U_2(j))$$
$$= \exp(U)\exp(U_1(i)) \sum_{j=1}^{J} \exp(U_2(j))$$

por el otro,

$$I_{ij} = U + U_1(i) + U_2(j) \Rightarrow p_{ij} = \exp(U + U_1(i) + U_2(j))$$
$$\Rightarrow p_{+j} = \sum_{i=1}^{I} \exp(U + U_1(i) + U_2(j)) = \sum_{i=1}^{I} \exp(U)\exp(U_1(i))\exp(U_2(j))$$
$$= \exp(U)\exp(U_2(j)) \sum_{i=1}^{I} \exp(U_1(i))$$

Además tenemos que

$$p_{++} = \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ij} = 1$$

es decir,

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \exp\left(U + U_1(i) + U_2(j)\right) = \sum_{i=1}^{I} \sum_{j=1}^{J} \exp(U) \exp\left(U_1(i)\right) \exp\left(U_2(j)\right)$$

$$= \exp(U) \sum_{i=1}^{I} \exp\left(U_1(i)\right) \sum_{j=1}^{J} \exp\left(U_2(j)\right) = 1$$

Finalmente, tenemos que el producto

$$p_{i+} \cdot p_{+j} = \frac{p_{i+} \cdot p_{+j}}{1} = \frac{p_{i+} \cdot p_{+j}}{p_{++}}$$

$$= \frac{\left[\exp(U) \exp\left(U_1(i)\right) \sum_{j=1}^{J} \exp\left(U_2(j)\right)\right] \cdot \left[\exp(U) \exp\left(U_2(j)\right) \sum_{i=1}^{I} \exp\left(U_1(i)\right)\right]}{\exp(U) \sum_{i=1}^{I} \exp\left(U_1(i)\right) \sum_{j=1}^{J} \exp\left(U_2(j)\right)}$$

$$= \exp(U) \exp\left(U_1(i)\right) \exp\left(U_2(j)\right) = \exp\left(U + U_1(i) + U_2(j)\right) = p_{ij}$$

q.e.d.

Resta probar que

$$p_{ij} = p_{i+} \cdot p_{+j} \Rightarrow I_{ij} = U + U_1(i) + U_2(j)$$

En el cuarto capítulo veremos que

$$U_{12}(i,j) = \frac{\sum_{r=1}^{I} \sum_{s=1}^{J} \ln\left(\theta_{ij,rs}\right)}{IJ} \quad \text{donde} \quad \theta_{ij,rs} = \frac{p_{ij} \cdot p_{rs}}{p_{rj} \cdot p_{is}},$$

entonces

$$p_{ij} = p_{i+} \cdot p_{+j} \Rightarrow \theta_{ij,rs} = \frac{\left(p_{i+} \cdot p_{+j}\right)\left(p_{r+} \cdot p_{+s}\right)}{\left(p_{r+} \cdot p_{+j}\right)\left(p_{i+} \cdot p_{+s}\right)} = 1$$

$$\Rightarrow U_{12}(i,j) = \frac{\sum_{r=1}^{I} \sum_{s=1}^{J} \ln\left(\theta_{ij,rs}\right)}{IJ} = \frac{\sum_{r=1}^{I} \sum_{s=1}^{J} \ln(1)}{IJ} = 0$$

q.e.d.

2.1.1 Bondad de ajuste y selección del modelo

Más adelante, en la sección de estimación, veremos que la estimación del conjunto $\{p_{ij}\}$ a partir del modelo saturado, es decir bajo la hipótesis de dependencia, es la siguiente:

$$\tilde{p}_{ij} = \frac{O_{ij}}{n}$$

ó, visto de otra manera,

$$\tilde{E}_{ij} = n \cdot \tilde{p}_{ij} = O_{ij}$$

Entonces tendríamos la siguiente tabla de valores esperados $\tilde{E}_{ij}$ en el estudio de asociación entre tipo de personalidad y filiación política:

	Demócrata	*Republicano*	*Suma*
Introvertido	30	70	100
Extrovertido	60	40	100
Suma	90	110	200

Esto nos lleva a

$$\chi^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\left(O_{ij} - \tilde{E}_{ij}\right)^2}{\tilde{E}_{ij}} = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(O_{ij} - O_{ij})^2}{O_{ij}} = 0.$$

Ahora, calculemos los grados de libertad (gl) con la fórmula

$$gl = \#parámetos\ nulos\ independientes$$

El modelo saturado incluye a todos los parámetros, por lo tanto ningún parámetro es nulo, y tenemos que $gl = 0$.

Debido a que $gl = 0$ no tiene sentido esta prueba y es necesario ajustar modelos no saturados en busca de la relación entre las variables bajo estudio.

Veremos más adelante que la estimación del conjunto $\{p_{ij}\}$ a partir del modelo $I_{ij} = U + U_1(i) + U_2(j)$, es decir bajo el supuesto de independencia, es la siguiente:

$$\tilde{p}_{ij} = \frac{O_{i+}}{n} \cdot \frac{O_{+j}}{n}$$

ó, visto de otra manera,

$$\tilde{E}_{ij} = n \cdot \tilde{p}_{ij} = \frac{O_{i+} \cdot O_{+j}}{n}$$

Entonces tendríamos la siguiente tabla de valores esperados $\tilde{E}_{ij}$ bajo la hipótesis de independencia entre tipo de personalidad y filiación política:

	Demócrata	*Republicano*	*Suma*
Introvertido	45	55	100
Extrovertido	45	55	100
Suma	90	110	200

Esto nos lleva al mismo valor de χ^2 que vimos en el capítulo anterior:

$$\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\left(O_{ij} - \tilde{E}_{ij}\right)^2}{\tilde{E}_{ij}} = 8.18.$$

En el modelo de independencia hay $I \times J$ parámetros nulos correspondientes a U_{12}, pero tan sólo se anulan $(I-1) \times (J-1)$ que son independientes pues $\sum_{i=1}^{I} U_{12}(i,j) = \sum_{j=1}^{J} U_{12}(i,j) = 0$.

Entonces

$$gl = \#\textit{parámetos nulos independientes} = (I-1) \times (J-1) = (2-1) \times (2-1) = 1.$$

Entonces la hipótesis de independencia entre tipo de personalidad y filiación política es inaceptable pues

$$\chi^2 = 18.18 > \chi^2_{.95} = 3.84 \; con\; gl = 1$$

Por lo tanto aceptamos que el modelo saturado

$$I_{ij} = U + U_1 + U_2 + U_{12}$$

es el que mejor describe la relación entre tipo de personalidad y filiación política y podemos concluir que existe dependencia entre estas variables.

2.1.2 Colapsabilidad

Ahora veamos un resultado que nos indicará cuando podemos colapsar las tablas de contingencia de dos dimensiones sin afectar la relación de las variables.

Las tablas independientes son colapsables pues a partir del modelo de independencia, tenemos que la diferencia entre logaritmos de cualquier par de celdas en la misma columna es la siguiente:

$$I_{ij} - I_{rj} = \ln(p_{ij}) - \ln(p_{rj}) = \ln\left(\frac{p_{ij}}{p_{rj}}\right) = U_1(i) - U_1(r)$$

y, cambiando la escala logarítmica a la escala original, tenemos que

$$\frac{p_{ij}}{p_{rj}} = \exp(U_1(i) - U_1(r)) = \frac{\exp(U_1(i))}{\exp(U_1(r))} = \frac{\exp(U)\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))}{\exp(U)\exp(U_1(r))\sum_{j=1}^{J}\exp(U_2(j))}$$

$$= \frac{\sum_{j=1}^{J}\exp(U)\exp(U_1(i))\exp(U_2(j))}{\sum_{j=1}^{J}\exp(U)\exp(U_1(r))\exp(U_2(j))} = \frac{\sum_{j=1}^{J}\exp(U+U_1(i)+U_2(j))}{\sum_{j=1}^{J}\exp(U+U_1(r)+U_2(j))}$$

$$= \frac{\sum_{j=1}^{J}\exp(\ln(p_{ij}))}{\sum_{j=1}^{J}\exp(\ln(p_{rj}))} = \frac{\sum_{j=1}^{J}p_{ij}}{\sum_{j=1}^{J}p_{rj}} = \frac{p_{i+}}{p_{r+}}$$

Así que el cociente entre celdas de la misma columna es equivalente al cociente de las correspondientes sumas por renglones. Esto también se cumple para elementos por renglon y las correspondientes sumas por columnas. Esto lleva a las siguientes dos conclusiones en las tablas independientes[2]:

i) la independencia no se pierde si combinamos las categorías de las variables.

ii) los parámetros U_1 pueden ser determinados a partir del vector de sumas por renglones $\{E_{i+}\}$ y análogamente los parámetros U_2 se pueden calcular a partir de las sumas por columnas.

Podemos combinar una o más categorías de alguna de las variables y el mismo modelo describe la estructura de la tabla colapsada. Inversamente, si la estructura no es independiente, combinar categorías nos lleva a una tabla con parámetros distintos de la tabla original; incluso los parámetros U_{12} podrían anularse, mostrando así una tabla colapsada con distinta estructura que la de la tabla original.

2.2 Modelos para tablas de tres dimensiones

Cuando tenemos tres variables categóricas, podemos ordenar los datos en una tabla tri-dimensional, la cual esta conformada por K tablas de contingencia bi-dimensionales, una por cada nivel de la tercera variable.

Podemos describir cada una de estas tablas bi-dimensionales con el modelo logarítmico-lineal que vimos en la sección anterior.

$I_{ijk}^{(k)} = V^{(k)} + V_1^{(k)}(i) + V_2^{(k)}(j) + V_{12}^{(k)}(i,j) \quad para \quad i = 1,\ldots,I \;\; y \;\; para \;\; j = 1,\ldots J \;\; y \;\; para \;\; k = 1,\ldots,K$

Para obtener los parámetros del modelo logarítmico-lineal para tablas de tres dimensiones podemos promediar estos parámetros y calcular las desviaciones de dichos promedios.

$$i) \quad U = \frac{\sum_{k=1}^{K} V^{(k)}}{K}$$

$$ii) \quad U_1(i) = \frac{\sum_{k=1}^{K} V_1^{(k)}(i)}{K}$$

$$iii) \quad U_2(j) = \frac{\sum_{k=1}^{K} V_2^{(k)}(j)}{K}$$

$$iv) \quad U_{12}(i,j) = \frac{\sum_{k=1}^{K} V_{12}^{(k)}(i,j)}{K}$$

$$v) \quad U_3(k) = V^{(k)} - U$$

$$vi) \quad U_{13}(i,k) = V_1^{(k)}(i) - U_1(i)$$

$$vii) \quad U_{23}(j,k) = V_2^{(k)}(j) - U_2(j)$$

$$viii) \quad U_{123}(i,j,k) = V_{12}^{(k)}(i,j) - U_{12}(i,j)$$

Los primeros cuatro parámetros provienen del promedio de los que suman cero en cada tabla bi-dimensional, esto lleva a que su suma también sea cero. Los restantes cuatro son desviaciones de un promedio y por lo tanto también suman cero. Por ejemplo,

$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

Ahora podemos escribir el modelo logarítmico-lineal saturado para la tabla $I \times J \times K$.

$$I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) + U_{12}(i,j) + U_{13}(i,k) + U_{23}(j,k) + U_{123}(i,j,k)$$

El efecto bi-factorial está definido como el promedio de la interacción entre las variables 1 y 2 en cada tabla de dos dimensiones, pero también podemos definir a U_{13} y U_{23} como promedios de la interacción en cada una de las tablas de la variable restante.

El efecto tri-factorial U_{123} cuantifica la diferencia de la magnitud del efecto bi-factorial U_{12} entre las tablas de dos dimensiones. Además, si cualquier efecto bi-factorial es constante a través de estas, tenemos que el efecto tri-factorial es cero. Esto nos lleva a la formulación del principio de jerarquía que veremos a continuación.

2.2.1 Modelos jerárquicos

Si tenemos dos parámetros, con r y s sub-índices respectivamente. Diremos que el parámetro con r sub-índices es de orden superior si r > s y además los r-subíndices de este contienen a los s subíndices del primer parámetro. Por ejemplo U_{123} es de orden superior a U_{12}, U_{13} y a U_{23}. Así como U_{12} es de orden superior a U_1 y a U_2.

La familia de modelos jerárquicos es aquella en la cual los parámetros de orden superior son nulos si los parámetros de orden inferior son nulos. E, inversamente, si un parámetro no es nulo entonces los parámetros de orden superior a este aparecen también en el modelo. Por ejemplo, si $U_{12} = 0$ tendremos que $U_{123} = 0$ y si U_{13} está presente en el modelo entonces U_1 y U_3 estarán presentes en el modelo también.

La mayoría de las tablas pueden ser descritas por modelos jerárquicos pero hay algunas excepciones. Además la interpretación de modelos no-jerárquicos es compleja. Bishop et. al[3]. presentan un modelo en el cual $U_{12} = U_{13} = U_{23} = 0$, pero $U_{123} \neq 0$. Este modelo está relacionado con el concepto de "sinergismo" pues el efecto se da cuando las tres variables están presentes pero no cuando están presentes tan sólo dos de estas.

2.2.2 Modelos comprensivos

Son aquellos que incluyen el efecto principal y los efectos de un sólo factor. Por ejemplo en el caso de tres variables los modelos comprensivos incluyen al menos los parámetros U, U_1, U_2 y U_3.

Inversamente, los modelos que no incluyen al menos estos parámetros son llamados no comprensivos.

2.2.3 Modelos no saturados, comprensivos y jerárquicos

Ahora veamos cómo los modelos no saturados corresponden a distintos tipos de relación entre las variables bajo estudio.

Cuando las tres variables son mutuamente independientes tenemos el modelo

$$I_{ijk} = U + U_1 + U_2 + U_3$$

que, como veremos más adelante, equivale a la siguiente aseveración

$\Pr(A_i \cap B_j \cap C_k) = \Pr(A_i)\Pr(B_j)\Pr(C_k)$ *para* $i = 1, \ldots, I$, *para* $j = 1, \ldots, J$ *y para* $k = 1, \ldots, K$.

donde A_i representa la i-ésima categoría de la primera variable, B_j la j-ésima categoría de la segunda variable y C_k la k-ésima categoría de la tercera variable.

Podemos escribir la igualdad anterior así:

$$p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k} \ \ para \ \ i = 1, \ldots, I \ \ , \ \ para \ \ j = 1, \ldots, J \ \ y \ para \ \ k = 1, \ldots, K$$

donde

$$\begin{aligned}
p_{ijk} &= \Pr\left(A_i \cap B_j \cap C_k\right) \\
p_{i++} &= \sum_{j=1}^{J} \sum_{k=1}^{K} p_{ijk} = \Pr\left(A_i\right) \\
p_{+j+} &= \sum_{i=1}^{I} \sum_{k=1}^{K} p_{ijk} = \Pr\left(B_j\right) \\
p_{++k} &= \sum_{i=1}^{I} \sum_{j=1}^{J} p_{ijk} = \Pr\left(C_k\right)
\end{aligned}$$

A continuación probaremos estas igualdades:

$$p_{i++} = \sum_{j=1}^{J} \sum_{k=1}^{K} p_{ijk} = \sum_{j=1}^{J} \sum_{k=1}^{K} \Pr\left(A_i \cap B_j \cap C_k\right) =$$

$$\sum_{j=1}^{J} \Pr\left(((A_i \cap B_j) \cap C_1) \cup ((A_i \cap B_j) \cap C_2) \cup \cdots \cup ((A_i \cap B_j) \cap C_K)\right) = \sum_{j=1}^{J} \Pr\left((A_i \cap B_j) \cap \left(\cup_{k=1}^{K} C_k\right)\right)$$

$$= \sum_{j=1}^{J} \Pr\left((A_i \cap B_j) \cap \Omega\right) = \sum_{j=1}^{J} \Pr\left(A_i \cap B_j\right) = \Pr\left((A_i \cap B_1) \cup (A_i \cap B_2) \cup \cdots \cup (A_i \cap B_J)\right)$$

$$= \Pr\left(A_i \cap \left(\cup_{j=1}^{J} B_j\right)\right) = \Pr\left(A_i \cap \Omega\right) = \Pr\left(A_i\right).$$

Cabe mencionar que la tercera igualdad se debe a que las categorías de cualquier variable categórica deben ser excluyentes, es decir, $C_t \cap C_k = \emptyset \ \ si \ \ t \neq k$. Lo anterior lleva a que $((A_i \cap B_j) \cap C_t) \cap ((A_i \cap B_j) \cap C_k) = \emptyset \ \ si \ \ t \neq k$ y por lo tanto a que $\Pr\left(((A_i \cap B_j) \cap C_t) \cap ((A_i \cap B_j) \cap C_k)\right) = \Pr\left((A_i \cap B_j) \cap C_t\right) + \Pr\left((A_i \cap B_j) \cap C_k\right) \ \ si \ \ t \neq k$.

Además, la quinta igualdad se debe a que las variables categóricas son exhaustivas, es decir, $\cup_{k=1}^{K} C_k = \Omega$.

Análogamente obtenemos que

$$p_{+j+} = \sum_{i=1}^{I}\sum_{k=1}^{K} p_{ijk} = \sum_{i=1}^{I}\sum_{k=1}^{K} \Pr(A_i \cap B_j \cap C_k)$$
$$= \sum_{i=1}^{I} \Pr(((A_i \cap B_j) \cap C_1) \cup ((A_i \cap B_j) \cap C_2) \cup \cdots \cup ((A_i \cap B_j) \cap C_K))$$
$$= \sum_{i=1}^{I} \Pr\left((A_i \cap B_j) \cap \left(\cup_{k=1}^{K} C_k\right)\right) = \sum_{i=1}^{I} \Pr((A_i \cap B_j) \cap \Omega)$$
$$= \sum_{i=1}^{I} \Pr(A_i \cap B_j) = \Pr((B_j \cap A_1) \cup (B_j \cap A_2) \cup \cdots \cup (B_j \cap A_I))$$
$$= \Pr\left(B_j \cap \left(\cup_{i=1}^{I} A_i\right)\right) = \Pr(B_j \cap \Omega) = \Pr(B_j)$$

y que

$$p_{++k} = \sum_{i=1}^{I}\sum_{j=1}^{J} p_{ijk} = \sum_{i=1}^{I}\sum_{j=1}^{J} \Pr(A_i \cap B_j \cap C_k)$$
$$= \sum_{i=1}^{I} \Pr(((A_i \cap C_k) \cap B_1) \cup ((A_i \cap C_k) \cap B_2) \cup \cdots \cup ((A_i \cap C_k) \cap B_J))$$
$$= \sum_{i=1}^{I} \Pr\left((A_i \cap C_k) \cap \left(\cup_{j=1}^{J} B_j\right)\right) = \sum_{i=1}^{I} \Pr((A_i \cap C_k) \cap \Omega) = \sum_{i=1}^{I} \Pr(A_i \cap C_k)$$
$$= \Pr((C_k \cap A_1) \cup (C_k \cap A_2) \cup \cdots \cup (C_k \cap A_I)) = \Pr\left(C_k \cap \left(\cup_{i=1}^{I} A_i\right)\right) = \Pr(C_k \cap \Omega) = \Pr(C_k)$$

q.e.d.

Ahora probaremos que

$$I_{ijk} = U + U_1 + U_2 + U_3 \Leftrightarrow p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k}.$$

Probemos primero que

$$I_{ijk} = U + U_1 + U_2 + U_3 \Rightarrow p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k}$$

Por un lado

$$I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) \Rightarrow p_{ijk} = \exp(U + U_1(i) + U_2(j) + U_3(k))$$

$$\Rightarrow p_{i++} = \sum_{j=1}^{J}\sum_{k=1}^{K} \exp(U + U_1(i) + U_2(j) + U_3(k)) = \sum_{j=1}^{J}\sum_{k=1}^{K} \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))$$
$$= \exp(U)\exp(U_1(i)) \sum_{j=1}^{J} \exp(U_2(j)) \sum_{k=1}^{K} \exp(U_3(k))$$

por el otro,

$$I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) \Rightarrow p_{ijk} = \exp(U + U_1(i) + U_2(j) + U_3(k))$$

$$\Rightarrow p_{+j+} = \sum_{i=1}^{I}\sum_{k=1}^{K} \exp(U + U_1(i) + U_2(j) + U_3(k)) = \sum_{i=1}^{I}\sum_{k=1}^{K} \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))$$

$$= \exp(U)\exp(U_2(j))\sum_{i=1}^{I}\exp(U_1(i))\sum_{k=1}^{K}\exp(U_3(k))$$

también

$$I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) \Rightarrow p_{ijk} = \exp(U + U_1(i) + U_2(j) + U_3(k))$$

$$\Rightarrow p_{++k} = \sum_{i=1}^{I}\sum_{j=1}^{J} \exp(U + U_1(i) + U_2(j) + U_3(k)) = \sum_{i=1}^{I}\sum_{j=1}^{J} \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))$$

$$= \exp(U)\exp(U_3(k))\sum_{i=1}^{I}\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))$$

Además $$p_{+++} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk} = 1$$

es decir,

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp(U + U_1(i) + U_2(j) + U_3(k)) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))$$

$$= \exp(U)\sum_{i=1}^{I}\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))\sum_{k=1}^{K}\exp(U_3(k)) = 1$$

Finalmente, tenemos que el producto

$$p_{i++}\cdot p_{+j+}\cdot p_{++k} = \frac{p_{i++}\cdot p_{+j+}\cdot p_{++k}}{1} = \frac{p_{i++}\cdot p_{+j+}}{p_{+++}}\cdot\frac{p_{++k}}{\cdot p_{+++}} =$$

$$\frac{\left[\exp(U)\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))\sum_{k=1}^{K}\exp(U_3(k))\right]\left[\exp(U)\exp(U_2(j))\sum_{i=1}^{I}\exp(U_1(i))\sum_{k=1}^{K}\exp(U_3(k)\right.}{\left[\exp(U)\sum_{i=1}^{I}\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))\sum_{k=1}^{K}\exp(U_3(k))\right]}$$

$$\cdot\frac{\left[\exp(U)\exp(U_3(k))\sum_{i=1}^{I}\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))\right]}{\left[\exp(U)\sum_{i=1}^{I}\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))\sum_{k=1}^{K}\exp(U_3(k))\right]} = \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))$$

$$= \exp(U + U_1(i) + U_2(j) + U_3(k)) = p_{ijk}$$

q.e.d.

Resta probar que

$$p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k} \Rightarrow I_{ijk} = U + U_1(i) + U_2(j) + U_3(k)$$

En el cuarto capítulo veremos que

$$U_{12}(i,j) = \frac{\sum\limits_{k=1}^{K}\sum\limits_{r=1}^{I}\sum\limits_{s=1}^{J} \ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK}$$
$$U_{13}(i,k) = \frac{\sum\limits_{j=1}^{J}\sum\limits_{r=1}^{I}\sum\limits_{t=1}^{K} \ln\left(\theta_{ik,rt}^{(j)}\right)}{IJK}$$
$$U_{23}(j,k) = \frac{\sum\limits_{i=1}^{I}\sum\limits_{s=1}^{J}\sum\limits_{t=1}^{K} \ln\left(\theta_{jk,st}^{(i)}\right)}{IJK}$$

donde

$$\theta_{ij,rs}^{(k)} = \frac{p_{ijk} \cdot p_{rsk}}{p_{rjk} \cdot p_{isk}}$$
$$\theta_{ik,rt}^{(j)} = \frac{p_{ijk} \cdot p_{rjt}}{p_{rjk} \cdot p_{ijt}}$$
$$\theta_{jk,st}^{(i)} = \frac{p_{ijk} \cdot p_{ist}}{p_{isk} \cdot p_{ijt}}.$$

Entonces

$$p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k} \Rightarrow \theta_{ij,rs}^{(k)} = \frac{(p_{i++} \cdot p_{+j+} \cdot p_{++k})(p_{r++} \cdot p_{+s+} \cdot p_{++k})}{(p_{r++} \cdot p_{+j+} \cdot p_{++k})(p_{i++} \cdot p_{+s+} \cdot p_{++k})} = 1 \Rightarrow U_{12}(i,j) =$$
$$\frac{\sum\limits_{k=1}^{K}\sum\limits_{r=1}^{I}\sum\limits_{s=1}^{J} \ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK} = \frac{\sum\limits_{k=1}^{K}\sum\limits_{r=1}^{I}\sum\limits_{s=1}^{J} \ln(1)}{IJK} = 0$$

$$p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k} \Rightarrow \theta_{ik,rt}^{(j)} = \frac{(p_{i++} \cdot p_{+j+} \cdot p_{++k})(p_{r++} \cdot p_{+j+} \cdot p_{++t})}{(p_{r++} \cdot p_{+j+} \cdot p_{++k})(p_{i++} \cdot p_{+j+} \cdot p_{++t})} = 1 \Rightarrow U_{13}(i,k) =$$
$$\frac{\sum\limits_{j=1}^{J}\sum\limits_{r=1}^{I}\sum\limits_{t=1}^{K} \ln\left(\theta_{ik,rt}^{(j)}\right)}{IJK} = \frac{\sum\limits_{j=1}^{J}\sum\limits_{r=1}^{I}\sum\limits_{t=1}^{K} \ln(1)}{IJK} = 0$$

$$p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k} \Rightarrow \theta_{jk,st}^{(i)} = \frac{(p_{i++} \cdot p_{+j+} \cdot p_{++k})(p_{i++} \cdot p_{+s+} \cdot p_{++t})}{(p_{i++} \cdot p_{+s+} \cdot p_{++k})(p_{i++} \cdot p_{+j+} \cdot p_{++t})} = 1 \Rightarrow U_{23}(j,k) =$$
$$\frac{\sum\limits_{i=1}^{I}\sum\limits_{s=1}^{J}\sum\limits_{t=1}^{K} \ln\left(\theta_{jk,st}^{(i)}\right)}{IJK} = \frac{\sum\limits_{i=1}^{I}\sum\limits_{s=1}^{J}\sum\limits_{t=1}^{K} \ln(1)}{IJK} = 0$$

También veremos en el cuarto capítulo que

$$V_{12}^{(k)}(i,j) = \frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJ}$$

lo cual nos lleva a que

$$p_{ijk} = p_{i++}\cdot p_{+j+}\cdot p_{++k} \Rightarrow \theta_{ij,rs}^{(k)} = \frac{(p_{i++}\cdot p_{+j+}\cdot p_{++k})(p_{r++}\cdot p_{+s+}\cdot p_{++k})}{(p_{r++}\cdot p_{+j+}\cdot p_{++k})(p_{i++}\cdot p_{+s+}\cdot p_{++k})} = 1 \Rightarrow V_{12}^{(k)}(i,j) =$$

$$\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJ} = \frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln(1)}{IJ} = 0$$

$$\Rightarrow U_{123}(i,j,k) = V_{12}^{(k)}(i,j) - U_{12}(i,j) = 0 - 0 = 0$$

q.e.d.

Hay tres versiones del modelo con sólo un efecto bifactorial. Si una variable, digamos la segunda, es conjuntamente independiente de las otras dos tenemos el modelo

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13}$$

que, como veremos más adelante, equivale a la siguiente aseveración

$\Pr(A_i \cap B_j \cap C_k) = \Pr(B_j)\Pr(A_i \cap C_k)$ *para* $i = 1,\ldots,I$, *para* $j = 1,\ldots,J$ *y para* $k = 1,\ldots,K$,

la cual también podemos escribir así:

$p_{ijk} = p_{+j+}\cdot p_{i+k}$ *para* $i = 1,\ldots,I$, *para* $j = 1,\ldots,J$ *y para* $k = 1,\ldots,K$

donde

$$p_{ijk} = \Pr(A_i \cap B_j \cap C_k)$$

$$p_{+j+} = \sum_{i=1}^{I}\sum_{k=1}^{K} p_{ijk} = \Pr(B_j)$$

$$p_{i+k} = \sum_{j=1}^{J} p_{ijk} = \Pr(A_i \cap C_k)$$

Probaremos la tercera igualdad que es la que aún no hemos probado:

$$p_{i+k} = \sum_{j=1}^{J} p_{ijk} = \sum_{j=1}^{J} \Pr\left(A_i \cap B_j \cap C_k\right) =$$
$$= \Pr\left(((A_i \cap C_k) \cap B_1) \cup ((A_i \cap C_k) \cap B_2) \cup \ldots \cup ((A_i \cap C_k) \cap B_J)\right) = \Pr\left((A_i \cap C_k) \cap \left(\cup_{j=1}^{J} B_j\right)\right)$$
$$= \Pr\left((A_i \cap C_k) \cap \Omega\right) = \Pr\left(A_i \cap C_k\right)$$

q.e.d.

A continuación probaremos que

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13} \Leftrightarrow p_{ijk} = p_{+j+} \cdot p_{i+k}.$$

Probemos primero que

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13} \Rightarrow p_{ijk} = p_{+j+} \cdot p_{i+k}.$$

Por un lado

$$I_{ijk} = U+U_1(i)+U_2(j)+U_3(k)+U_{13}(i,k) \Rightarrow p_{ijk} = \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{123}(i,k)\right)$$
$$\Rightarrow p_{+j+} = \sum_{i=1}^{I}\sum_{k=1}^{K} \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k)\right)$$
$$= \sum_{i=1}^{I}\sum_{k=1}^{K} \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))\exp(U_{13}(i,k))$$
$$= \exp(U)\exp(U_2(j))\sum_{i=1}^{I}\sum_{k=1}^{K} \exp(U_1(i))\exp(U_3(k))\exp(U_{13}(i,k))$$

por el otro

$$I_{ijk} = U+U_1(i)+U_2(j)+U_3(k)+U_{13}(i,k) \Rightarrow p_{ijk} = \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{123}(i,k)\right)$$
$$\Rightarrow p_{i+k} = \sum_{j=1}^{J} \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k)\right)$$
$$= \sum_{j=1}^{J} \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))\exp(U_{13}(i,k))$$
$$= \exp(U)\exp(U_1(i))\exp(U_3(k))\exp(U_{13}(i,k))\sum_{j=1}^{J}\exp(U_2(j)).$$

Además $$p_{+++} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} p_{ijk} = 1$$

es decir,

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k)\right)$$
$$= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))\exp(U_{13}(i,k))$$
$$= \exp(U) \sum_{j=1}^{J} \exp(U_2(j)) \sum_{i=1}^{I} \sum_{k=1}^{K} \exp(U_1(i))\exp(U_3(k))\exp(U_{13}(i,k)) = 1$$

Finalmente, tenemos que el producto

$$p_{+j+} \cdot p_{i+k} = \frac{p_{+j+} \cdot p_{i+k}}{1} = \frac{p_{+j+}}{p_{+++}} \cdot p_{i+k}$$
$$= \frac{\left[\exp(U)\exp(U_2(j)) \sum_{i=1}^{I} \sum_{k=1}^{K} \exp(U_1(i))\exp(U_3(k))\exp(U_{13}(i,k))\right]}{\exp(U) \sum_{j=1}^{J} \exp(U_2(j)) \sum_{i=1}^{I} \sum_{k=1}^{K} \exp(U_1(i))\exp(U_3(k))\exp(U_{13}(i,k))}$$
$$\cdot \left[\exp(U)\exp(U_1(i))\exp(U_3(k))\exp(U_{13}(i,k)) \sum_{j=1}^{J} \exp(U_2(j))\right]$$
$$= \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))\exp(U_{13}(i,k))$$
$$= \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k)\right) = p_{ijk}$$

q.e.d.

Resta probar que

$$p_{ijk} = p_{+j+} \cdot p_{i+k} \Rightarrow I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k).$$

$$p_{ijk} = p_{+j+} \cdot p_{i+k} \Rightarrow \theta_{ij,rs}^{(k)} = \frac{(p_{+j+} \cdot p_{i+k})(p_{+s+} \cdot p_{r+k})}{(p_{+j+} \cdot p_{r+k})(p_{+s+} \cdot p_{i+k})} = 1$$
$$\Rightarrow U_{12}(i,j) = \frac{\sum_{k=1}^{K} \sum_{r=1}^{I} \sum_{s=1}^{J} \ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK} = \frac{\sum_{k=1}^{K} \sum_{r=1}^{I} \sum_{s=1}^{J} \ln(1)}{IJK} = 0$$

$$p_{ijk}=p_{+j+}\cdot p_{i+k}\Rightarrow \theta_{jk,st}^{(i)}=\frac{(p_{+j+}\cdot p_{i+k})(p_{+s+}\cdot p_{i+t})}{(p_{+s+}\cdot p_{i+k})(p_{+j+}\cdot p_{i+t})}=1\Rightarrow U_{23}(j,k)=\frac{\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\ln\left(\theta_{jk,st}^{(i)}\right)}{IJK}=$$
$$\frac{\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\ln(1)}{IJK}=0$$

$$p_{ijk}=p_{+j+}\cdot p_{i+k}\Rightarrow \theta_{ij,rs}^{(k)}=\frac{(p_{+j+}\cdot p_{i+k})(p_{+s+}\cdot p_{r+k})}{(p_{+j+}\cdot p_{r+k})(p_{+s+}\cdot p_{i+k})}=1\Rightarrow V_{12}^{(k)}(i,j)=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJ}=$$
$$\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln(1)}{IJ}=0$$
$$\Rightarrow U_{123}(i,j,k)=V_{12}^{(k)}(i,j)-U_{12}(i,j)=0-0=0$$

q.e.d.

Hay tres versiones del modelo sin un efecto bi-factorial. Si dos variables, digamos la 1 y la 2 son independientes en todos los niveles de la otra variable, en este caso la tercera, decimos que son condicionalmente independientes y tenemos el modelo

$$l_{ijk}=U+U_1+U_2+U_3+U_{13}+U_{23}$$

que, como veremos más adelante, equivale a la siguiente aseveración

$\Pr(A_i\cap B_j\mid C_k)=\dfrac{\Pr(A_i\cap C_j)\Pr(B_j\cap C_k)}{\Pr(C_k)}$ *para* $i=1,\ldots,I$, *para* $j=1,\ldots,J$ *y para* $k=1,\ldots,K$,

la cual también podemos escribir así:

$$p_{ijk}=\frac{p_{i+k}\cdot p_{+jk}}{p_{++k}}\quad para\ \ i=1,\ldots,I\ ,\ \ para\ \ j=1,\ldots,J\ \ y\ \ para\ \ k=1,\ldots,K$$

donde

$$p_{ijk}=\Pr(A_i\cap B_j\cap C_k)$$
$$p_{i+k}=\sum_{j=1}^{J}p_{ijk}=\Pr(A_i\cap C_k)$$
$$p_{+jk}=\sum_{i=1}^{I}p_{ijk}=\Pr(B_j\cap C_k)$$
$$p_{++k}=\sum_{i=1}^{I}\sum_{j=1}^{J}p_{ijk}=\Pr(C_k)$$

Probaremos la tercera igualdad que es la que aún no hemos probado:

$$p_{+jk} = \sum_{i=1}^{I} p_{ijk} = \sum_{i=1}^{I} \Pr(A_i \cap B_j \cap C_k) =$$

$$\Pr(((B_j \cap C_k) \cap A_1) \cup ((B_j \cap C_k) \cap A_2) \cup \ldots \cup ((B_j \cap C_k) \cap A_I))$$

$$= \Pr\left((B_j \cap C_k) \cap \left(\cup_{i=1}^{I} A_i\right)\right) = \Pr((B_j \cap C_k) \cap \Omega) = \Pr(B_j \cap C_k)$$

q.e.d.

A continuación probaremos que

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13} + U_{23} \Leftrightarrow p_{ijk} = \frac{p_{i+k} \cdot p_{+jk}}{p_{++k}}.$$

Probemos primero que

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13} + U_{23} \Rightarrow p_{ijk} = \frac{p_{i+k} \cdot p_{+jk}}{p_{++k}}.$$

Por un lado

$$I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k)$$

$$\Rightarrow p_{ijk} = \exp(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k))$$

$$\Rightarrow p_{i+k} = \sum_{j=1}^{J} \exp(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k))$$

$$= \sum_{j=1}^{J} \exp(U) \exp(U_1(i)) \exp(U_2(j)) \exp(U_3(k)) \exp(U_{13}(i,k)) \exp(U_{23}(j,k))$$

$$= \exp(U) \exp(U_1(i)) \exp(U_3(k)) \exp(U_{13}(i,k)) \sum_{j=1}^{J} \exp(U_2(j)) \exp(U_{23}(j,k)).$$

por el otro

$$I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k)$$

$$\Rightarrow p_{ijk} = \exp(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k))$$

$$\Rightarrow p_{+jk} = \sum_{i=1}^{I} \exp(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k))$$

$$= \sum_{i=1}^{I} \exp(U) \exp(U_1(i)) \exp(U_2(j)) \exp(U_3(k)) \exp(U_{13}(i,k)) \exp(U_{23}(j,k))$$

$$= \exp(U) \exp(U_2(j)) \exp(U_3(k)) \exp(U_{23}(i,k)) \sum_{i=1}^{I} \exp(U_1(i)) \exp(U_{13}(i,k)).$$

Además

$$I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k)$$

$$\Rightarrow p_{ijk} = \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k)\right)$$

$$\Rightarrow p_{++k} = \sum_{i=1}^{I}\sum_{j=1}^{J} \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k)\right)$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J} \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))\exp(U_{13}(i,k))\exp(U_{23}(j,k))$$

$$= \exp(U)\exp(U_3(k))\sum_{i=1}^{I}\sum_{j=1}^{J} \exp(U_1(i))\exp(U_2(j))\exp(U_{13}(i,k))\exp(U_{23}(i,k)).$$

Finalmente, tenemos que el cociente

$$\frac{p_{i+k}}{p_{++k}} \cdot p_{+jk} =$$

$$\frac{\left[\exp(U)\exp(U_1(i))\exp(U_3(k))\exp(U_{13}(i,k))\sum_{j=1}^{J}\exp(U_2(j))\exp(U_{23}(j,k))\right]}{\exp(U)\exp(U_3(k))\sum_{i=1}^{I}\sum_{j=1}^{J}\exp(U_1(i))\exp(U_2(j))\exp(U_{13}(i,k))\exp(U_{23}(i,k))}$$

$$\cdot\left[\exp(U)\exp(U_2(j))\exp(U_3(k))\exp(U_{23}(i,k))\sum_{i=1}^{I}\exp(U_1(i))\exp(U_{13}(i,k))\right]$$

$$= \exp(U)\exp(U_1(i))\exp(U_2(j))\exp(U_3(k))\exp(U_{13}(i,k))\exp(U_{23}(j,k))$$

$$= \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k)\right) = p_{ijk}$$

q.e.d.

Resta probar que

$$p_{ijk} = \frac{p_{i+k}\cdot p_{+jk}}{p_{++k}} \Rightarrow I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k).$$

$$p_{ijk} = \frac{p_{i+k}\cdot p_{+jk}}{p_{++k}} \Rightarrow \theta_{ij,rs}^{(k)} = \frac{\left[\left(\frac{p_{i+k}\cdot p_{+jk}}{p_{++k}}\right)\left(\frac{p_{r+k}\cdot p_{+sk}}{p_{++k}}\right)\right]}{\left[\left(\frac{p_{r+k}\cdot p_{+jk}}{p_{++k}}\right)\left(\frac{p_{i+k}\cdot p_{+sk}}{p_{++k}}\right)\right]} = 1$$

$$\Rightarrow U_{12}(i,j) = \frac{\sum_{k=1}^{K}\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK} = \frac{\sum_{k=1}^{K}\sum_{r=1}^{I}\sum_{s=1}^{J}\ln(1)}{IJK} = 0$$

$$p_{ijk}=\frac{p_{i+k}\cdot p_{+jk}}{p_{++k}}\Rightarrow\theta^{(k)}_{ij,rs}=\frac{\left[\left(\frac{p_{i+k}\cdot p_{+jk}}{p_{++k}}\right)\left(\frac{p_{r+k}\cdot p_{+sk}}{p_{++k}}\right)\right]}{\left[\left(\frac{p_{r+k}\cdot p_{+jk}}{p_{++k}}\right)\left(\frac{p_{i+k}\cdot p_{+sk}}{p_{++k}}\right)\right]}=1$$

$\Rightarrow V^{(k)}_{12}(i,j)=\dfrac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta^{(k)}_{ij,rs}\right)}{IJ}=\dfrac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln(1)}{IJ}=0\Rightarrow U_{123}(i,j,k)=V^{(k)}_{12}(i,j)-U_{12}(i,j)=$ $0-0=0$

q.e.d.

Cuando no varía la asociación parcial entre las variables tenemos el modelo

$$l_{ijk}=U+U_1+U_2+U_3+U_{12}+U_{13}+U_{23}\,.$$

En este caso no existe una expresión "directa" para p_{ijk} como en los casos anteriores, pero podemos ajustar este modelo mediante el método de "ajuste proporcional iterativo" que veremos más adelante en la sección de estimación.

2.2.4 Colapsabilidad

Obtuvimos el modelo para tres variables al promediar a través de las tablas parciales. Ahora veremos cuándo podemos obtener los parámetros, sin caer en equivocaciones, a partir de la tabla marginal.

Teorema[4]:

Una variable es colapsable en una tabla de tres dimensiones si y sólo si es condicionalmente independiente de una de las otras variables dada la tercera.

Veamos a continuación un ejercicio para ejemplificar la colapsabilidad.

Los datos de la siguiente tabla, analizada por Bishop en 1969[5], representa la sobrevivencia de los infantes (variable 1) según el cuidado prenatal recibido por parte de sus madres (variable 2). Los cuidados son clasificados en "mucho" y "poco". Las madres asistieron a dos clínicas, que denotaremos como A y B.

	Cuidado Prenatal	Sobrevivencia Sí	No	Tasa de mortalidad
Clínica A	Poco	3	176	1.7 %
	Mucho	4	293	1.4 %
Clínica B	Poco	17	197	7.9 %
	Mucho	2	23	8.0 %

En la Clínica A tenemos el cociente de ventajas muestral es

$$\tilde{\theta}^{(1)} = \frac{(3)(293)}{(4)(176)} = 1.2$$

y en la Clínica B

$$\tilde{\theta}^{(2)} = \frac{(17)(23)}{(2)(197)} = 1.0.$$

Debido a que ambos son aproximadamente 1 podemos concluir que $U_{12} = 0$, es decir, que el cuidado prenatal no está asociado con la sobrevivencia del infante.

Pero si colapsamos respecto a la variable 3 (Clínica) tendríamos siguiente tabla marginal.

Cuidado Prenatal	Sobrevivencia Sí	No
Poco	20	373
Mucho	6	316

Ahora tenemos que el cociente de ventajas muestral es

$$\tilde{\theta} = \frac{(20)(316)}{(6)(373)} = 2.8$$

que no refleja la verdadera relación entre cuidado prenatal y sobrevivencia. Cometeríamos un error si analizáramos esta asociación a partir de la tabla marginal. Esto sucede, según el teorema anterior, porque U_{13} y U_{23} son ambos distintos de cero. Es decir que si alguno de estos fuese zero entonces podríamos colapsar a través de la variable 3 y no se verían afectados los parámetros del modelo.

2.3 Modelos para tablas de cuatro o más dimensiones

Procedimos de dos a tres dimensiones al modelar cada tabla bi-dimensional, una por cada categoría de la tercera variable. Análogamente, cuando tenemos un arreglo de cuatro dimensiones, podemos modelar cada arreglo tri-dimensional en cada uno de las L categorías de la cuarta variable. Podemos continuar con este proceso para cualquier número de dimensiones, digamos S[6].

La siguiente tabla muestra la relación entre los parámetros W para tres dimensiones y los parámetros U para cuatro dimensiones.

		Parámetros U	
Orden	*Parámetros W para cada tabla tri − dimensional*	*Promedio de parámetros W*	*Desviación del promedio**
Promedio	W	U	U_4
Uni − factorial	W_1, W_2, W_3	U_1, U_2, U_3	U_{14}, U_{24}, U_{34}
Bi − factorial	W_{12}, W_{23}, W_{13}	U_{12}, U_{23}, U_{13}	$U_{124}, U_{234}, U_{134}$
Tri − factorial	W_{123}	U_{123}	U_{1234}

* Notemos que las desviaciones son de orden mayor que los demás términos de cada renglón.

REFERENCIAS

[1]Bishop, Yvonne M.M. et. al., Discrete Multivariate Analysis, págs. 16-17

[2]Ibidem., pág. 29

[3]Ibidem., pág. 38

[4]Ibidem., pág. 39

[5]Ibidem., pág. 41

[6]Ibidem., págs. 42-43

CAPÍTULO 3

ESTIMACIÓN Y BONDAD DE AJUSTE

3.1 Conjunto mínimo de estadísticas suficientes

Antes de seleccionar un modelo, utilizaremos la información muestral contenida en la tabla de contingencia para hallar los estimadores máximo-verosímiles (EMV) de las frecuencias esperadas en las celdas de la tabla de contingencia. Estos dependen únicamente de los datos a través del conjunto mínimo de estadísticas suficientes, el cual contiene toda la información que aporta la muestra para el conocimiento de los parámetros que deseamos estimar[1].

Primero veremos cómo se obtiene dicho conjunto con el caso de tres variables para después generalizarlo a cualquier número de variables.

Para facilitar la exposición, asumiremos el modelo de muestreo más simple: las observaciones muestrales $\{O_{ijk}\}$ de la tabla de contingencia son variables aleatorias independientes tipo Poisson con valores esperados $\{E_{ijk}\}$, donde $E_{ijk} = E(O_{ijk})$.

Llamaremos función de verosimilitud al producto de las funciones de probabilidad de cada una de las variables aleatorias[2], también llamada función de probabilidad conjunta.

La función de verosimilitud Poisson de $\{O_{ijk}\}$ es

$$L(\{O_{ijk}\},\{E_{ijk}\}) = \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\frac{\exp(-E_{ijk})\cdot E_{ijk}^{O_{ijk}}}{O_{ijk}!}.$$

Este esquema es el adecuado cuando de antemano es desconocida la cantidad de individuos que serán muestreados.

Los EMV de $\{E_{ijk}\}$ son aquellos que, al ser sustituidos en la función de verosimilitud, la llevan a alcanzar su máximo valor posible. Para hallarlos es suficiente maximizar $\ln(L)$ pues la función $\ln(x)$ con $x > 0$ es estrictamente creciente[3].

$$\begin{aligned}\ln(L(\{O_{ijk}\},\{E_{ijk}\})) &= \ln\left(\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\frac{\exp(-E_{ijk})\cdot E_{ijk}^{O_{ijk}}}{O_{ijk}!}\right) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln\left(\frac{\exp(-E_{ijk})\cdot E_{ijk}^{O_{ijk}}}{O_{ijk}!}\right)\\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln(\exp(-E_{ijk})) + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln\left(E_{ijk}^{O_{ijk}}\right) - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln(O_{ijk}!)\\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(-E_{ijk}) + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\ln(E_{ijk}) - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln(O_{ijk}!)\end{aligned}$$

El tercer sumando es constante pues $\{O_{ijk}\}$ son las observaciones muestrales en la tabla de contingencia. Entonces debe maximizarse

$$K(\{O_{ijk}\},\{E_{ijk}\}) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\ln(E_{ijk}) - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} E_{ijk}$$

que es llamado el kernel, o núcleo, del logaritmo de la función de verosimilitud.

Ahora, consideremos el modelo para tablas de tres dimensiones:

$$l_{ijk} = \ln(p_{ijk}) = U + U_1(i) + U_2(j) + U_3(k) + U_{12}(i,j) + U_{13}(i,k) + U_{23}(j,k) + U_{123}(i,j,k)$$

Si definimos $l_{ijk} = \ln(E_{ijk})$ en vez de $l_{ijk} = \ln(p_{ijk})$ la misma ecuación es aplicable[4]. Entonces tenemos que

$$\begin{aligned}
&K(\{O_{ijk}\},\{E_{ijk}\}) \\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\left(U + U_1(i) + U_2(j) + U_3(k) + U_{12}(i,j) + U_{13}(i,k) + U_{23}(j,k) + U_{123}(i,j,k)\right) \\
&- \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{12}(i,j) + U_{13}(i,k) + U_{23}(j,k) + U_{123}(i,j,k)\right) \\
&= n\cdot U + \sum_{i=1}^{I} O_{i++}\cdot U_1(i) + \sum_{j=1}^{J} O_{+j+}\cdot U_2(j) + \sum_{k=1}^{K} O_{++k}\cdot U_3(k) + \sum_{i=1}^{I}\sum_{j=1}^{J} O_{ij+}\cdot U_{12}(i,j) \\
&+ \sum_{i=1}^{I}\sum_{k=1}^{K} O_{i+k}\cdot U_{13}(i,k) + \sum_{j=1}^{J}\sum_{k=1}^{K} O_{+jk}\cdot U_{23}(j,k) + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\cdot U_{123}(i,j,k) \\
&- \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{12}(i,j) + U_{13}(i,k) + U_{23}(j,k) + U_{123}(i,j,k)\right)
\end{aligned}$$

donde

$$n = O_{+++} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\ ,\quad O_{i++} = \sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\ ,\quad O_{+j+} = \sum_{i=1}^{I}\sum_{k=1}^{K} O_{ijk}\ ,$$

$$O_{++k} = \sum_{i=1}^{I}\sum_{j=1}^{J} O_{ijk}\ ,\quad O_{ij+} = \sum_{k=1}^{K} O_{ijk}\ ,\quad O_{i+k} = \sum_{j=1}^{J} O_{ijk}\ \ y\ \ O_{+jk} = \sum_{i=1}^{I} O_{ijk}.$$

Podemos reescribir la función de verosimilitud Poisson de la siguiente manera:

$$L(\{O_{ijk}\},\{E_{ijk}\}) = \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\frac{\exp(-E_{ijk})\cdot E_{ijk}^{O_{ijk}}}{O_{ijk}!} = \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\frac{\exp(-E_{ijk})\cdot\exp(O_{ijk}\cdot\ln(E_{ijk}))}{O_{ijk}!}$$

$$= \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\frac{\exp(O_{ijk}\cdot\ln(E_{ijk})-E_{ijk})}{O_{ijk}!} = \frac{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\exp(-E_{ijk}+O_{ijk}\cdot\ln(E_{ijk}))}{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}!}$$

$$= \frac{\exp\left(\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(O_{ijk}\cdot\ln(E_{ijk})-E_{ijk})\right)}{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}!} = a(\{E_{ijk}\})\,b(\{O_{ijk}\})\exp\left(\sum_{j=1}^{IJK}c_j(\{E_{ijk}\})\,d_j(\{O_{ijk}\})\right)$$

donde

$$a(\{E_{ijk}\}) = 1$$

$$b(\{O_{ijk}\}) = \frac{1}{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}!}$$

$$c_j(\{E_{ijk}\}) = (O_j\cdot\ln(E_j)-E_j)\quad para\quad j=1,\ldots,IJK$$

$$d_j(\{O_{ijk}\}) = 1\ \ para\ \ j=1,\ldots,IJK$$

Lo cual muestra que la distribución conjunta Poisson pertenece a la familia exponencial con k-parámetros[5], entonces los coeficientes de los parámetros son estadísticas suficientes[6]. También Bishop et. al. corroboran esto: "las estadísticas suficientes de valores reales son los términos adyacentes a los parámetros desconocidos"[7].

También podemos argumentar la suficiencia de estas estadísticas con el Teorema Fischer-Neyman ó Criterio de Factorización para estadísticas conjuntamente suficientes[8]:

"Sea X_1, X_2, ... ,X_n una muestra aleatoria de la función de probabilidad f (X ; θ), donde el parámetro θ puede ser un vector. Un conjunto de estadísticas $S_1 = s_1(X_1, X_2, \ldots ,X_n)$, ... , $S_r = s_r(X_1, X_2, \ldots ,X_n)$ es conjuntamente suficiente si y sólo si la función de probabilidad conjunta conjunta de $\{X_i\}$ puede ser factorizada así:

$$f\left(\{X_i\}\ ;\ \theta\right) = g\left(s_1(X_1,\ X_2,\ \ldots, X_n),\ \ldots,\ s_r(X_1,\ X_2,\ \ldots, X_n)\, ;\theta\right) \cdot h\left(\{X_i\}\right)$$

donde $h(\{X_i\})$ es no negativa y no involucra al parámetro θ y la función $g(s_1, \ldots, s_r\ ;\ \theta)$ depende de $\{X_i\}$ sólo a través de las funciones $s_1(\bullet, \ldots, \bullet), \ldots,\ S_r = s_r(\bullet, \ldots, \bullet)$."

En nuestro caso tenemos

$$h\left(\{O_{ijk}\}\right) = \frac{1}{\prod_{i=1}^{I} \prod_{j=1}^{J} \prod_{k=1}^{K} O_{ijk}!}$$

la cual es no negativa, en particular positiva (bajo el supuesto $O_{ijk} > 0$), y no depende del parámetro $\theta = (E_{111}, \ldots, E_{IJK})$ y además

$$g\left(\{O_{ijk}\}, \{E_{ijk}\}\right) = \exp\left(K\left(\{O_{ijk}\}, \{E_{ijk}\}\right)\right)$$

que depende de $\{O_{ijk}\}$ sólo a través de

$$S_1\left(\{O_{ijk}\}\right) = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} O_{ijk} = O_{+++} = n$$

$$S_2\left(\{O_{ijk}\}\right) = \sum_{j=1}^{J} \sum_{k=1}^{K} O_{ijk} = O_{i++} \quad para\ \ i = 1, \ldots, I$$

$$S_3\left(\{O_{ijk}\}\right) = \sum_{i=1}^{I} \sum_{k=1}^{K} O_{ijk} = O_{+j+} \quad para\ \ j = 1, \ldots, J$$

$$S_4\left(\{O_{ijk}\}\right) = \sum_{i=1}^{I} \sum_{j=1}^{J} O_{ijk} = O_{++k} \quad para\ \ k = 1, \ldots, K$$

$$S_5\left(\{O_{ijk}\}\right) = \sum_{k=1}^{K} O_{ijk} = O_{ij+} \quad para\ \ i = 1, \ldots, I\ \ y\ \ para\ \ j = 1, \ldots, J$$

$$S_6\left(\{O_{ijk}\}\right) = \sum_{j=1}^{J} O_{ijk} = O_{i+k} \quad para\ \ i = 1, \ldots, I\ \ y\ \ para\ \ k = 1, \ldots, K$$

$$S_7\left(\{O_{ijk}\}\right) = \sum_{i=1}^{I} O_{ijk} = O_{+jk} \quad para\ \ j = 1, \ldots, J\ \ y\ \ para\ \ k = 1, \ldots, K$$

$$S_8\left(\{O_{ijk}\}\right) = O_{ijk} \quad para\ \ i = 1, \ldots, I\ ,\ \ para\ \ j = 1, \ldots, J\ \ y\ \ para\ \ k = 1, \ldots, K$$

Por lo tanto estas son estadísticas conjuntamente suficientes y por lo tanto contienen toda la información que aporta la muestra para el conocimiento del parámetro $\theta = (E_{111}, \ldots, E_{IJK})$.

Algunas de estas estadísticas son redundantes pues pueden ser calculadas a partir de otras, por lo tanto podemos resumir la información muestral todavía más en un conjunto más pequeño de estadísticas suficientes que llamaremos conjunto mínimo de estadísticas suficientes.

En el caso del modelo saturado el Kernel del logaritmo de la función de verosimilitud permanece igual:

$$
\begin{aligned}
&K\left(\{O_{ijk}\},\{E_{ijk}\}\right) \\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\left(U+U_1(i)+U_2(j)+U_3(k)+U_{12}(i,j)+U_{13}(i,k)+U_{23}(j,k)+U_{123}(i,j,k)\right) \\
&- \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp\left(U+U_1(i)+U_2(j)+U_3(k)+U_{12}(i,j)+U_{13}(i,k)+U_{23}(j,k)+U_{123}(i,j,k)\right) \\
&\quad = n\cdot U+\sum_{i=1}^{I} O_{i++}\cdot U_1(i)+\sum_{j=1}^{J} O_{+j+}\cdot U_2(j)+\sum_{k=1}^{K} O_{++k}\cdot U_3(k)+\sum_{i=1}^{I}\sum_{j=1}^{J} O_{ij+}\cdot U_{12}(i,j) \\
&\qquad + \sum_{i=1}^{I}\sum_{k=1}^{K} O_{i+k}\cdot U_{13}(i,k)+\sum_{j=1}^{J}\sum_{k=1}^{K} O_{+jk}\cdot U_{23}(j,k)+\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\cdot U_{123}(i,j,k) \\
&- \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp\left(U+U_1(i)+U_2(j)+U_3(k)+U_{12}(i,j)+U_{13}(i,k)+U_{23}(j,k)+U_{123}(i,j,k)\right)
\end{aligned}
$$

y tenemos que el conjunto mínimo de estadísticas suficientes es

$$O_{ijk} \;\; para \;\; i=1,\ldots,I \;\; , \;\; para \;\; j=1,\ldots,J \;\; y \;\; para \;\; k=1,\ldots,K$$

pues podemos calcular las restantes estadísticas suficientes a partir de estas. Cabe mencionar que a esto se debe que la estimación del modelo saturado dé un ajuste $\chi^2=0$.

Pero si tenemos un modelo no saturado, por ejemplo:

$$I_{ijk}=U+U_1+U_2+U_3+U_{12}+U_{13}+U_{23}$$

el Kernel se simplifica como vemos a continuación pues $U_{123}=0$:

$$
\begin{aligned}
&K\left(\{O_{ijk}\},\{E_{ijk}\}\right) \\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\left(U+U_1(i)+U_2(j)+U_3(k)+U_{12}(i,j)+U_{13}(i,k)+U_{23}(j,k)+U_{123}(i,j,k)\right) \\
&- \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp\left(U+U_1(i)+U_2(j)+U_3(k)+U_{12}(i,j)+U_{13}(i,k)+U_{23}(j,k)+U_{123}(i,j,k)\right) \\
&\quad = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} O_{ijk}\left(U+U_1(i)+U_2(j)+U_3(k)+U_{12}(i,j)+U_{13}(i,k)+U_{23}(j,k)\right) \\
&\quad - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp\left(U+U_1(i)+U_2(j)+U_3(k)+U_{12}(i,j)+U_{13}(i,k)+U_{23}(j,k)\right) \\
&\quad = n\cdot U+\sum_{i=1}^{I} O_{i++}\cdot U_1(i)+\sum_{j=1}^{J} O_{+j+}\cdot U_2(j)+\sum_{k=1}^{K} O_{++k}\cdot U_3(k)+\sum_{i=1}^{I}\sum_{j=1}^{J} O_{ij+}\cdot U_{12}(i,j)
\end{aligned}
$$

$$+\sum_{i=1}^{I}\sum_{k=1}^{K} O_{i+k}\cdot U_{13}(i,k)+\sum_{j=1}^{J}\sum_{k=1}^{K} O_{+jk}\cdot U_{23}(j,k)$$
$$-\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\exp\left(U+U_1(i)+U_2(j)+U_3(k)+U_{12}(i,j)+U_{13}(i,k)+U_{23}(j,k)\right)$$

y tenemos que el conjunto mínimo de estadísticas suficientes es

$$O_{ij+}=\sum_{k=1}^{K} O_{ijk} \quad para\ i=1,\ldots,I\ y\ para\ j=1,\ldots,J$$
$$O_{i+k}=\sum_{j=1}^{J} O_{ijk} \quad para\ i=1,\ldots,I\ y\ para\ k=1,\ldots,K$$
$$O_{+jk}=\sum_{i=1}^{I} O_{ijk} \quad para\ j=1,\ldots,J\ y\ para\ k=1,\ldots,K$$

pues podemos calcular las restantes estadísticas suficientes a partir de estas.

Podemos resumir el proceso de obtención del conjunto mínimo de estadísticas suficientes en la siguiente tabla, la cual muestra los conjuntos mínimos de estadísticas suficientes correspondientes a los diversos modelos jerárquicos y comprensivos para tres variables categóricas:

Modelo jerárquico y comprensivo	*Estadísticas suficientes mínimas*
$I_{ijk}=U+U_1+U_2+U_3$	$\{O_{i++}\},\{O_{+j+}\},\{O_{++k}\}$
$I_{ijk}=U+U_1+U_2+U_3+U_{12}$	$\{O_{ij+}\},\{O_{++k}\}$
$I_{ijk}=U+U_1+U_2+U_3+U_{13}$	$\{O_{i+k}\},\{O_{+j+}\}$
$I_{ijk}=U+U_1+U_2+U_3+U_{23}$	$\{O_{+jk}\},\{O_{i++}\}$
$I_{ijk}=U+U_1+U_2+U_3+U_{12}+U_{13}$	$\{O_{ij+}\},\{O_{i+k}\}$
$I_{ijk}=U+U_1+U_2+U_3+U_{12}+U_{23}$	$\{O_{ij+}\},\{O_{+jk}\}$
$I_{ijk}=U+U_1+U_2+U_3+U_{13}+U_{23}$	$\{O_{i+k}\},\{O_{+jk}\}$
$I_{ijk}=U+U_1+U_2+U_3+U_{12}+U_{13}+U_{23}$	$\{O_{ij+}\},\{O_{i+k}\},\{O_{+jk}\}$
$I_{ijk}=U+U_1+U_2+U_3+U_{12}+U_{13}+U_{23}+U_{123}$	$\{O_{ijk}\}$

Ahora pasemos a la generalización del proceso de obtención del conjunto mínimo de estadísticas suficientes, para ello denotemos con θ al conjunto de todos los posibles sub-índices de los parámetros del modelo formados con elementos de $S=\{1,2,\ldots,s\}$, donde s es el número de variables categóricas cuya relación se está analizando.

En realidad tenemos que θ es el conjunto potencia de S y que los sub-índices de los parámetros de cualquier modelo, que son 2^S en el caso del modelo saturado, son elementos de θ.

La tabla marginal que llamaremos correspondiente al parámetro U_{θ_i} se obtiene al sumar los elementos de la tabla de contingencia a través de las variables que no están incluidas en el sub-índice θ_i.

Finalmente tenemos que el conjunto mínimo de estadísticas suficientes consiste en las tablas marginales correspondientes a los parámetros de orden superior incluidos en el modelo y, en el caso en que los sub-índices de estos no abarcan todas las variables, el conjunto mínimo de estadísticas suficientes también incluye a las tablas marginales correspondientes a los parámetros de orden inmediato inferior y así sucesivamente hasta que los sub-índices lleguen a abarcar a todas las variables.

3.2 Resultados de Birch

A las frecuencias esperadas en las tablas marginales resultantes del procedimiento que recién vimos se les conoce como configuraciones suficientes. Birch[9] demostró que los elementos de dichas tablas marginales son los EMV de los elementos que conforman las configuraciones suficientes. Es decir, si tenemos el sub-índice $\theta - \theta_i$, donde $\theta_i \in \theta$, entonces

$$\sum_{\theta_i} O_{\theta_S}$$

es el EMV de los elementos en la configuración $C_{\theta-\theta_i}$ donde $\theta_S \in \theta$ es el sub-índice que incluye a todas las variables, es decir, que O_{θ_S} es elemento de la tabla de contingencia.

Para ilustrar el primero de los resultados de Birch que acabamos de enunciar nos valdremos del módelo de independencia condicional entre las variables 1 y 2:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13} + U_{23}.$$

En este modelo, el kernel del logaritmo de la función de verosimilitud Poisson es:

$$K(\{O_{ijk}\}, \{E_{ijk}\})$$

$$= n \cdot U + \sum_{i=1}^{I} O_{i++} \cdot U_1(i) + \sum_{j=1}^{J} O_{+j+} \cdot U_2(j) + \sum_{k=1}^{K} O_{++k} \cdot U_3(k) + \sum_{i=1}^{I}\sum_{k=1}^{K} O_{i+k} \cdot U_{13}(i,k)$$
$$+ \sum_{j=1}^{J}\sum_{k=1}^{K} O_{+jk} \cdot U_{23}(j,k) - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \exp\left(U + U_1(i) + U_2(j) + U_3(k) + U_{13}(i,k) + U_{23}(j,k)\right).$$

Ahora veamos las derivadas parciales del Kernel respecto a los parámetros del modelo que nos ocupa en este ejemplo:

$$\frac{\delta K}{\delta U} = n - E_{+++}$$
$$\frac{\delta K}{\delta U_1(i)} = O_{i++} - E_{i++}$$
$$\frac{\delta K}{\delta U_2(j)} = O_{+j+} - E_{+j+}$$
$$\frac{\delta K}{\delta U_3(k)} = O_{++k} - E_{++k}$$
$$\frac{\delta K}{\delta U_{13}(i,k)} = O_{i+k} - E_{i+k}$$
$$\frac{\delta K}{\delta U_{23}(j,k)} = O_{+jk} - E_{+jk}.$$

Los valores con que se anulan las derivadas parciales son:

$$\tilde{E}_{+++} = n$$
$$\tilde{E}_{i++} = O_{i++}$$
$$\tilde{E}_{+j+} = O_{+j+}$$
$$\tilde{E}_{++k} = O_{++k}$$
$$\tilde{E}_{i+k} = O_{i+k}$$
$$\tilde{E}_{+jk} = O_{+jk}$$

y por lo tanto son los EMV de los valores en las celdas de las configuraciones suficientes.

El segundo resultado de Birch es la unicidad de la solución $\left\{\tilde{E}_{0s}\right\}$ que satisface tanto las restricciones del modelo como las estadísticas suficientes mínimas. Esto implica que si producimos una solución que iguale a las estadísticas suficientes mínimas, esta corresponde a los EMV de las frecuencias esperadas en la tabla de contingencia[10].

Cabe mencionar que Birch también demostró que los EMV de $\{E_{\theta s}\}$ son los mismos tanto para la función de verosimilitud Poisson como para la función de verosimilitud multinomial de $\{O_{\theta s}\}$[11] que es la adecuada cuando el tamaño de la muestra y el tamaño de la población son de antemano conocidos y que podemos ver a continuación[12]:

$$L\left(\{O_{ijk}\},\{E_{ijk}\}\right)=\frac{n!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}}\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\left(\frac{E_{ijk}}{N}\right)^{O_{ijk}}$$

donde n es el tamaño de la muestra y N es el tamaño de la población.

Esto se debe a que también forma parte de la familia de las funciones exponenciales y a que el kernel del logaritmo de la función de verosimilitud multinomial es muy parecido al de la función Poisson, tal como vemos a continuación:

$$L\left(\{O_{ijk}\},\{E_{ijk}\}\right)=$$

$$\frac{n!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}}\exp\left(\ln\left(\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\left(\frac{E_{ijk}}{N}\right)^{O_{ijk}}\right)\right)=\frac{n!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}}\exp\left(\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\ln\left(\frac{E_{ijk}}{N}\right)\right)$$

$$=a\left(\{p_{ijk}\}\right)b\left(\{O_{ijk}\}\right)\exp\left(\sum_{j=1}^{IJK}c_j\left(\{p_{ijk}\}\right)d_j\left(\{O_{ijk}\}\right)\right)$$

donde

$$a\left(\{p_{ijk}\}\right)=1$$

$$b\left(\{O_{ijk}\}\right)=\frac{n!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}!}$$

$$c_j\left(\{E_{ijk}\}\right)=\ln\left(\frac{E_j}{N}\right)\quad para\ \ j=1,\ldots,IJK$$

$$d_j\left(\{O_{ijk}\}\right)=O_j\ \ para\ \ j=1,\ldots,IJK$$

lo cual muestra que la función de verosimilitud multinomial forma parte de la familia exponencial con k-parámetros, además tenemos que:

$$\ln\left(L\left(\{O_{ijk}\},\{E_{ijk}\}\right)\right)=$$

$$\ln\left(\frac{n!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}}\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\left(\frac{E_{ijk}}{N}\right)^{O_{ijk}}\right)=\ln(n!)+\ln\left(\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}\left(\frac{E_{ijk}}{N}\right)^{O_{ijk}}\right)-\ln\left(\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K}O_{ijk}\right)$$

$$=\ln(n!)+\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\ln\left(\frac{E_{ijk}}{N}\right)-\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln\left(O_{ijk}\right)$$

$$=\ln(n!)+\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\left(\ln\left(E_{ijk}\right)-\ln(N)\right)-\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln\left(O_{ijk}\right)$$

$$=\ln(n!)+\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\ln\left(E_{ijk}\right)-\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\ln(N)-\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln\left(O_{ijk}\right)$$

$$=\ln(n!)+\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\ln\left(E_{ijk}\right)-\ln(N)\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}-\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln\left(O_{ijk}\right)$$

$$=\ln(n!)+\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\ln\left(E_{ijk}\right)-n\ln(N)-\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\ln\left(O_{ijk}\right),$$

lo cual muestra que el kernel del logaritmo de la función de verosimilitud multinomial es:

$$K\left(\{O_{ijk}\},\{E_{ijk}\}\right)=\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}O_{ijk}\ln\left(p_{ijk}\right)$$

pues el primero, tercero y cuarto sumandos son constantes.

3.3 Estimación

Para estimar los parámetros de un modelo jerárquico se consideran dos métodos diferentes[13], dependiendo de las características del modelo particular y que se presentan en forma resumida a continuación.

1er. método. Obtención de los EMV en forma "directa".

i) Al definir el modelo que se quiere ajustar se obtiene el conjunto mínimo de estadísticas suficientes.

ii) Se obtienen valores esperados en las celdas de cada configuración suficiente mediante el conjunto mínimo de estadísticas suficientes.

iii) Usando los resultados de Birch se obtienen los estimadores de la frecuencia esperada en cada celda de la tabla de contingencia.

2o. método. Ajuste proporcional iterativo

i) Se aplica el primer paso del método anteriormente descrito y se asignan los valores iniciales $\left\{\tilde{E}_{\theta s}^{(0)}\right\}$, generalmente:

$$\tilde{E}_{\theta s}^{(0)} = 1$$

ii) Para cada configuración suficiente se ajustan sucesivamente los estimadores preliminares en forma recursiva del siguiente modo:

Supongamos que el modelo para el cual se realiza la estimación consta de k configuraciones suficientes, a saber $C_{\theta_1}, C_{\theta_2}, \ldots, C_{\theta_k}$, con frecuencias esperadas $E_{\theta_1}, E_{\theta_2}, \ldots, E_{\theta_k}$ respectivamente.

$$\tilde{E}_{\theta s}^{(1)} = \tilde{E}_{\theta s}^{(0)} \frac{O_{\theta_1}}{\tilde{E}_{\theta_1}^{(0)}}$$
$$\tilde{E}_{\theta s}^{(2)} = \tilde{E}_{\theta s}^{(1)} \frac{O_{\theta_2}}{\tilde{E}_{\theta_2}^{(1)}}$$
$$\vdots$$
$$\tilde{E}_{\theta s}^{(k)} = \tilde{E}_{\theta s}^{(k-1)} \frac{O_{\theta_k}}{\tilde{E}_{\theta_k}^{(k-1)}}$$

Al término de esta primera etapa se le denomina ciclo, mismo que se repite hasta que los estimadores convergen con la precisión deseada, es decir hasta que:

$$\left|\tilde{E}_{\theta s}^{(kr)} - \tilde{E}_{\theta s}^{(k(r-1))}\right| < \delta$$

donde δ y r representan la precisión deseada y el ciclo respectivamente.

Si definimos:

$$G_{\theta s}^{2\,(k)} = -2\sum_{\theta s} O_{\theta s} \ln\left(\frac{\tilde{E}_{\theta s}}{O_{\theta s}}\right),$$

que es la medida de bondad de ajuste que veremos más adelante, tenemos que $G^2 \geq 0$[14], es decir que está acotada inferiormente, y es monótonamente decreciente conforme el ciclo avanza[15], por lo tanto converge al ínfimo.

Este método genera un conjunto de estimadores de los valores esperados en la tabla de contingencia que satisface tanto el modelo como las estadísticas suficientes, el cual es único según los resultados de Birch y además son EMV[16].

Para distinguir los modelos según el método de estimación que se pueden utilizar los dos siguientes teoremas[17]:

" Si el conjunto mínimo de estadísticas suficientes para un modelo consiste en dos configuraciones únicamente, entonces existen estimadores directos."

"Si cada una de las tres configuraciones tiene conjuntos diferentes de variables en común, respecto a las otras dos configuraciones, los estimadores directos de las celdas elementales no existen. Inversamente, la estimación directa es posible con tres configuraciones si:

i) Al menos dos no tienen variables en común, o

ii) Existe un conjunto de variables común a todas, y si este es eliminado, se aplica el inciso anterior."

Pero en realidad no es indispensable hacer esta distinción pues el "ajuste proporcional iterativo" converge al terminar el primer ciclo cuando las variables no pasan de ser seis[18].

Ahora veamos, con fines ilustrativos, el proceso de obtención de los EMV con el caso de tres variables.

En el capítulo anterior obtuvimos las expresiones para p_{ijk} equivalentes a los diversos modelos jerárquicos y comprensivos para tres variables, las cuales podemos resumir en el siguiente cuadro:

Modelo*	p_{ijk}	Tipo de independencia
(1, 2, 3)	$p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k}$	*A, B y C son conjuntamente independientes*
(12, 3)	$p_{ijk} = p_{++k} \cdot p_{ij+}$	*C es conjuntamente independiente de A y B*
(13, 2)	$p_{ijk} = p_{+j+} \cdot p_{i+k}$	*B es conjuntamente independiente de A y C*
(23, 1)	$p_{ijk} = p_{i++} \cdot p_{+jk}$	*A es conjuntamente independiente de B y C*
(12, 13)	$p_{ijk} = \dfrac{p_{ij+} \cdot p_{i+k}}{p_{+i+}}$	*B y C son condicionalmente independientes de A*
(12, 23)	$p_{ijk} = \dfrac{p_{ij+} \cdot p_{+jk}}{p_{+j+}}$	*A y C son condicionalmente independientes de B*
(13, 23)	$p_{ijk} = \dfrac{p_{i+k} \cdot p_{+jk}}{p_{++k}}$	*A y B son condicionalmente independientes de C*
(12, 13, 23)	*No existe expresión directa*	U_{12}, U_{13} *y* U_{23} *son constantes*
(123)	$p_{ijk} = p_{ijk}$	U_{12}, U_{13} *y* U_{23} *no son constantes*

* La notación empleada indica los sub-índices de las configuraciones suficientes correspondientes a cada modelo.

Si el tamaño de la muestra es conocido, digamos n, tenemos que

$$E_{ijk} = E\left(O_{ijk}\right) = E\left(n \cdot p_{ijk}\right) = n \cdot E\left(p_{ijk}\right) = n \cdot p_{ijk}.$$

Además los EMV de funciones de parámetros son simplemente las funciones evaluadas en los EMV de los parámetros[19]. En este caso tendríamos la siguiente lista de EMV:

$$\tilde{p}_{ijk} = \frac{\tilde{E}_{ijk}}{n} = \frac{O_{ijk}}{n}$$

$$\tilde{p}_{i++} = \sum_{j=1}^{J}\sum_{k=1}^{K} \tilde{p}_{ijk} = \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{\tilde{E}_{ijk}}{n} = \frac{\tilde{E}_{i++}}{n} = \frac{O_{i++}}{n}$$

$$\tilde{p}_{+j+} = \sum_{i=1}^{I}\sum_{k=1}^{K} \tilde{p}_{ijk} = \sum_{i=1}^{I}\sum_{k=1}^{K} \frac{\tilde{E}_{ijk}}{n} = \frac{\tilde{E}_{+j+}}{n} = \frac{O_{+j+}}{n}$$

$$\tilde{p}_{++k} = \sum_{i=1}^{I}\sum_{j=1}^{J} \tilde{p}_{ijk} = \sum_{i=1}^{I}\sum_{j=1}^{J} \frac{\tilde{E}_{ijk}}{n} = \frac{\tilde{E}_{++k}}{n} = \frac{O_{++k}}{n}$$

$$\tilde{p}_{ij+} = \sum_{k=1}^{K} \tilde{p}_{ijk} = \sum_{k=1}^{K} \frac{\tilde{E}_{ijk}}{n} = \frac{\tilde{E}_{ij+}}{n} = \frac{O_{ij+}}{n}$$

$$\tilde{p}_{i+k} = \sum_{j=1}^{J} \tilde{p}_{ijk} = \sum_{j=1}^{J} \frac{\tilde{E}_{ijk}}{n} = \frac{\tilde{E}_{i+k}}{n} = \frac{O_{i+k}}{n}$$

$$\tilde{p}_{+jk} = \sum_{i=1}^{I} \tilde{p}_{ijk} = \sum_{i=1}^{I} \frac{\tilde{E}_{ijk}}{n} = \frac{\tilde{E}_{+jk}}{n} = \frac{O_{+jk}}{n}$$

Finalmente podemos resumir en el siguiente cuadro los EMV para el estudio de tres variables categóricas:

Modelo	$\tilde{p}_{ijk}$
(1, 2, 3)	$\tilde{p}_{ijk} = \tilde{p}_{i++} \cdot \tilde{p}_{+j+} \cdot \tilde{p}_{++k} = \dfrac{O_{i++} \cdot O_{+j+} \cdot O_{++k}}{n^3}$
(12, 3)	$\tilde{p}_{ijk} = \tilde{p}_{++k} \cdot \tilde{p}_{ij+} = \dfrac{O_{++k} \cdot O_{ij+}}{n^2}$
(13, 2)	$\tilde{p}_{ijk} = \tilde{p}_{+j+} \cdot \tilde{p}_{i+k} = \dfrac{O_{+j+} \cdot O_{i+k}}{n^2}$
(23, 1)	$\tilde{p}_{ijk} = \tilde{p}_{i++} \cdot \tilde{p}_{+jk} = \dfrac{O_{i++} \cdot O_{+jk}}{n^2}$
(12, 13)	$\tilde{p}_{ijk} = \dfrac{\tilde{p}_{ij+} \cdot \tilde{p}_{i+k}}{\tilde{p}_{i++}} = \dfrac{O_{ij+} \cdot O_{i+k}}{n \cdot O_{i++}}$
(12, 23)	$\tilde{p}_{ijk} = \dfrac{\tilde{p}_{ij+} \cdot \tilde{p}_{+jk}}{\tilde{p}_{+j+}} = \dfrac{O_{ij+} \cdot O_{+jk}}{n \cdot O_{+j+}}$
(13, 23)	$\tilde{p}_{ijk} = \dfrac{\tilde{p}_{i+k} \cdot \tilde{p}_{+jk}}{\tilde{p}_{++k}} = \dfrac{O_{i+k} \cdot O_{+jk}}{n \cdot O_{++k}}$
(12, 13, 23)	*No existe expresión directa, se hará un ajuste proporcional iterativo*
(123)	$\tilde{p}_{ijk} = \dfrac{O_{ijk}}{n}$

No existe expresión "directa" para p_{ijk} cuando tenemos el modelo

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{12} + U_{13} + U_{23}.$$

Entonces es necesario realizar un ajuste proporcional iterativo cuyo primer ciclo vemos a continuación para hallar los EMV de las frecuencias esperadas en la tabla de contingencia correspondientes a este modelo.

$$\tilde{E}^{(1)}_{ijk} = \tilde{E}^{(0)}_{ijk} \frac{O_{ij+}}{\tilde{E}^{(0)}_{ij+}}$$
$$\tilde{E}^{(2)}_{ijk} = \tilde{E}^{(1)}_{ijk} \frac{O_{i+k}}{\tilde{E}^{(1)}_{i+k}}$$
$$\tilde{E}^{(3)}_{ijk} = \tilde{E}^{(3)}_{ijk} \frac{O_{+jk}}{\tilde{E}^{(3)}_{+jk}}.$$

Para ejemplificar el proceso de estimación, en el apéndice A se encuentra detallado el proceso de obtención de los EMV correspondientes a los diversos modelos para el estudio de asociación entre raza del defendido ó acusado, raza de la víctima y pena de muerte que comentamos en el primer capítulo.

3.4 Bondad de ajuste y selección del modelo

Ya que obtuvimos los EMV de las frecuencias esperadas en las tablas correspondientes a cada modelo, las podemos comparar con la tabla de contingencia. Usaremos estadísticas con distribución χ^2 para probar si las frecuencias esperadas en la tabla de contingencia satisfacen cierto modelo[20], las cuales son llamadas "pruebas de bondad de ajuste"[21].

Las estadísticas apropiadas para probar la bondad de ajuste de un modelo dado, llamémosle A, son las χ^2 de Pearson y el denominado cociente de máxima-verosimilitud G^2, las cuales definimos a continuación[22]:

$$\chi^{2(A)} = \sum_{\theta s} \frac{\left(O_{\theta s} - \tilde{E}_{\theta s}^{(A)}\right)^2}{\tilde{E}_{\theta s}^{(A)}}$$

$$G^{2(A)} = -2 \sum_{\theta s} O_{\theta s} \ln \left(\frac{\tilde{E}_{\theta s}^{(A)}}{O_{\theta s}} \right)$$

donde $\left\{ \tilde{E}_{\theta s}^{(A)} \right\}$ son los EMV de las frecuencias esperadas en la tabla de contingencia correspondientes al modelo A.

Ambas presentan asintóticamente, es decir conforme el tamaño de la muestra es mayor, distribución χ^2 con grados de libertad[23]:

$$gl = \textit{número de parámetros independientes nulos del modelo saturado.}$$

En el apéndice B se encuentra detallado el proceso de obtención de gl para el estudio de asociación entre tres variables categóricas.

Ahora veamos algunas propiedades del cociente de máxima-verosimilitud G^2 por las cuales lo utilizaremos preferentemente para medir la bondad de ajuste.

Los valores grandes de G^2 corresponden a probabilidades pequeñas en la cola derecha de la función de densidad χ^2 que llamaremos valores-p e indican que el ajuste es pobre[24].

Podemos reescribir G^2 de la siguiente manera:

$$G^{2(A)} = -2\sum_{\theta s} O_{\theta s}\left(\ln\left(\tilde{E}_{\theta s}^{(A)}\right) - \ln\left(O_{\theta s}\right)\right) = -2\left(\sum_{\theta s} O_{\theta s}\ln\left(\tilde{E}_{\theta s}^{(A)}\right) - \sum_{\theta s} O_{\theta s}\ln\left(O_{\theta s}\right)\right).$$

Si nos valemos del kernel del logaritmo de la función de verosimilitud multinomial:

$$K\left(\{E_{\theta s}\}\right) = \sum_{\theta s} O_{\theta s}\ln\left(E_{\theta s}\right)$$

tenemos que

$$G^{2(A)} = -2\left(K\left(\left\{\tilde{E}_{\theta s}^{(A)}\right\}\right) - K\left(\{O_{\theta s}\}\right)\right).$$

Si tenemos el modelo saturado, entonces

$$\tilde{E}_{\theta s} = O_{\theta s}$$

lo cual nos lleva a que

$$K\left(\left\{\tilde{E}_{\theta s}\right\}\right) = K\left(\{O_{\theta s}\}\right)$$

y por lo tanto a que

$$G^2 = -2\left(K\left(\{O_{\theta s}\}\right) - K\left(\{O_{\theta s}\}\right)\right) = 0.$$

Si tenemos un modelo A no saturado tenemos que

$$E_{\theta s}^{(A)} > 0$$

$$\sum_{\theta s}\tilde{E}_{\theta s}^{(A)} = \sum_{\theta s} O_{\theta s} = n$$

y por lo tanto que[25]

$$K\left(\left\{\tilde{E}_{\theta s}^{(A)}\right\}\right) < K\left(\{O_{\theta s}\}\right)$$

es decir que

$$G^{2(A)} = -2\left(K\left(\left\{\tilde{E}_{\theta s}^{(A)}\right\}\right) - K\left(\{O_{\theta s}\}\right)\right) > 0$$

y como los EMV se obtienen maximizando el kernel de la función de verosimilitud tenemos que G^2 es un mínimo[26].

Si tenemos un modelo A que contiene los parámetros de otro modelo, digamos B, tenemos que[27]:

$$G^{2(B/A)} = G^{2(B)} - G^{2(A)}$$

presenta distribución χ^2 con $V_2 - V_1$ grados de libertad, donde V_1 y V_2 son los grados de libertad correspondientes al modelo A y al modelo B respectivamente.

Podemos probar la significancia de cada uno de los parámetros del modelo saturado mediante $G^{2(B/A)}$, donde los modelos A y B difieren únicamente por el parámetro en cuestión. Hay que tener en cuenta que los valores-p pequeños correspondientes a $G^{2(B/A)}$ indican que el párametro es significativo.

Finalmente, para seleccionar el modelo tenemos que considerar que el valor-p del correspondiente cociente de verosimilitud G^2 tiene que ser mayor que el nivel de significancia α y además que el ajuste "interno", es decir en las celdas de la tabla de contingencia, es también satisfactorio. El ajuste interno puede ser medido con la desviación Freeman-Tukey[28]:

$$z_{ijk} = \sqrt{O_{ijk}} + \sqrt{O_{ijk} + 1} - \sqrt{4 \cdot \tilde{E}_{ijk} + 1}$$

los cuales se distribuyen $N(0,1)^{29}$.

Para ejemplificar la selección del modelo Agresti presenta el siguiente cuadro[30] para el estudio de asociación entre pena de muerte, raza del acusado y raza de la víctima, cuya obtención se encuentra detallada en el apéndice C:

Modelo	G^2	*gl*	*valor* − *p*
(1, 2, 3)	137.93	4	0.000
(12, 3)	8.13	3	0.043
(13, 2)	137.71	3	0.000
(23, 1)	131.68	3	0.000
(12, 13)	7.91	2	0.019
(12, 23)	1.88	2	0.390
(13, 23)	131.46	2	0.000
(12, 13, 23)	0.70	1	0.402
(123)	0	0	1.000

Esta tabla indica que los modelos $(1,2,3)$, $(13,2)$, $(23,1)$ *y* $(13,23)$ presentan un ajuste pobre. La característica que tienen en común estos modelos es la ausencia del parámetro U_{12} lo cual sugiere un asociación significativa entre raza del acusado y raza de la víctima. Podemos probar estadísticamente la sinificancia de U_{12} mediante la diferencia entre los cocientes de máxima verosimilitud de dos modelos que difieran únicamente por este parámetro, por ejemplo $A = (13,23)$ *y* $B = (12,13,23)$ nos llevan a que

$$G^{2(B/A)} = G^{2(B)} - G^{2(A)} = 131.46 - 0.70 = 130.76$$

$$gl = 2 - 1 = 1$$

$$valor - p = 0.000$$

lo cual nos indica que el parámetro U_{12} es significativo.

De los modelos que incluyen al parámetro U_{12} sólo los modelos $(12,23)$ *y* $(12,13,23)$ presentan valores-p mayores que $\alpha = 0.05$ de los cuales seleccionamos el más simple, a saber

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{12} + U_{23}.$$

Ahora, para seleccionar finalmente este modelo, hay que verificar el ajuste "interno":

Raza del acusado	*Raza de la víctima*	*Pena de muerte* *Sí*	*No*
Blanca	*Blanca*	−0.4249	0.2110
	Negra	−0.7113	0.2401
Negra	*Blanca*	0.7535	− 0.2625
	Negra	0.2920	− 0.0236

Debido a que todas las desviaciones Freeman-Tukey son menores en valor absoluto que $Z_{0.975} = 1.96$, tenemos que el ajuste interno es aceptable y por lo tanto el modelo que mejor representa la relación entre raza del acusado, raza de la víctima y pena de muerte es

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{12} + U_{23}$$

y por lo tanto podemos concluir que la raza del acusado y la raza de la víctima están asociadas, así como la raza de la víctima y la pena de muerte. Para conocer el sentido de ambas asociaciones tenemos que observar los valores de los EMV de los cocientes de ventajas como vemos a continuación:

$$\tilde{\theta}_{23}^{(i=1)}=\frac{21.1682\cdot 8.5179}{0.4821\cdot 129.8318}=2.88 \;\; y \;\; \tilde{\theta}_{23}^{(i=2)}=\frac{8.8318\cdot 97.4821}{5.5179\cdot 54.1682}=2.88$$

indican que tanto para los acusados de raza blanca ($i=1$) como para los acusados de raza negra ($i=2$) la aplicación de la pena de muerte es más frecuente cuando la víctima es de raza blanca,

$$\tilde{\theta}_{12}^{(k=1)}=\frac{21.1682\cdot 5.5179}{8.8318\cdot 0.4821}=27.43 \;\; y \;\; \tilde{\theta}_{12}^{(k=2)}=\frac{129.8318\cdot 97.4821}{54.1682\cdot 8.5179}=27.43$$

indican que tanto en los casos en que se aplicó pena de muerte ($k=1$) como en los que no se aplicó esta ($k=2$) , los acusados de raza blanca están relacionados con víctimas de raza blanca,

$$\tilde{\theta}_{13}^{(j=1)}=\frac{21.1682\cdot 54.1682}{8.8318\cdot 129.8318}=1.00 \;\; y \;\; \tilde{\theta}_{13}^{(j=2)}=\frac{0.4821\cdot 97.4821}{5.5179\cdot 8.5179}=1.00$$

indican que tanto en los casos en que la víctima es de raza blanca ($j=1$) como en los casos en que la víctima es de raza negra ($j=2$) la aplicación de la pena de muerte no está asociada con la raza del acusado.

REFERENCIAS

[1]Agresti, Alan, Categorical Data Analysis, págs. 165-166

[2]Mood, Alexander M. et. al., Introducción a la teoría de la estadística, pág. 208, citado por Villa Palato, María Soledad, Análisis estadístico de la población escolar de primer ingreso a la ENEP Acatlán a través de modelos log-lineales, pág. 36

[3]Roussas, George G., A first course in mathematical Statistics, págs. 242-243, citado por Villa Palato, María Soledad, op. cit., pág. 39

[4]Bishop, Yvonne M.M. et. al., Discrete Multivariate Analysis, pág. 19

[5]Mood, Alexander M. et. al., Introduction to the theory of statistics, pág. 313

[6]Agresti, Alan, op. cit., pág. 166

[7]Bishop, Yvonne M.M. et. al., op. cit., pág. 65, citado por Villa Palato, María Soledad, op. cit., pág. 40

[8]Mood, Alexander M. et. al., Introduction to the theory of statistics, págs. 307-308

[9]Agresti, Alan, op. cit., pág. 168

[10]Ibidem., pág. 169

[11]Ibidem., pág. 169

[12]Lehmann, E.L., Theory of point estimation, pág. 28, citado por Villa Palato, María Soledad, op. cit., págs. 38-39

[13]Bishop, Yvonne M.M. et. al., op. cit., pág. 65, citado por Villa Palato, María Soledad, op. cit., págs. 44-45

[14]Bishop, Yvonne M.M. et. al., op. cit., págs. 85-86, citado por Villa Palato, María Soledad, op. cit., págs. 45-46

[15]Agresti, Alan, op. cit., pág. 186

[16]Ibidem., pág. 186

[17]Bishop, Yvonne M.M. et. al., op. cit., págs. 78-79, citado por Villa Palato, María Soledad, op. cit., pág. 42

[18]Agresti, Alan, op. cit., pág. 186

[19]Mood, Alexander M. et. al., Introduction to the theory of statistics, pág. 285

[20]Agresti, Alan, op. cit., pág. 174

[21]Roussas, George G., op. cit., pág. 307, citado por Villa Palato, María Soledad, op. cit., pág. 47

[22]Villa Palato, María Soledad, op. cit., págs. 47-48

[23]Agresti, Alan, op. cit., pág. 174

[24]Ibidem., pág. 176

[25]Bishop, Yvonne M.M. et. al., op. cit., pág. 126, citado por Villa Palato, María Soledad, op. cit., pág. 49

[26]Villa Palato, María Soledad, op. cit., pág. 49

[27]Bishop, Yvonne M.M. et. al., op. cit., pág. 126, citado por Villa Palato, María Soledad, op. cit., págs. 49-50

[28]Bishop, Yvonne M.M. et. al., Discrete Multivariate Analysis, pág. 137

[29]Ibidem., pág. 137

[30]Agresti, Alan, op. cit., pág. 176

CAPÍTULO 4

INTERPRETACIÓN DE LOS PARÁMETROS DE ASOCIACIÓN

4.0 Interpretación de los parámetros de asociación del modelo logarítmico-lineal para dos variables categóricas

En el segundo capítulo definimos los parámetros de asociación, así nombrados por Agresti[1], de la siguiente manera:
$U_{12}(i,j) = I_{ij} - (U + U_1(i) + U_2(j))$ *para* $i = 1, \ldots, I$ *y para* $j = 1, \ldots, J$
donde

$$I_{ij} = \ln(p_{ij}) \text{ para } i = 1, \ldots, I \text{ y para } j = 1, \ldots, J$$

$$U = \frac{I_{++}}{IJ} \text{ donde } I_{++} = \sum_{i=1}^{I} \sum_{j=1}^{J} I_{ij}$$

$$U_1(i) = \frac{I_{i+}}{J} - \frac{I_{++}}{IJ} \text{ donde } I_{i+} = \sum_{j=1}^{J} I_{ij}$$

$$U_2(j) = \frac{I_{+j}}{I} - \frac{I_{++}}{IJ} \text{ donde } I_{+j} = \sum_{i=1}^{I} I_{ij}.$$

Para explicar el significado de los parámetros de asociación tanto Agresti[2] como Bishop et. al.[3] lo señalan como la desviación del supuesto de independencia pues, según demostramos en el segundo capítulo,

$$I_{ij} = U + U_1(i) + U_2(j) \text{ para } i = 1, \ldots, I \text{ y para } j = 1, \ldots, J$$

equivale a $p_{ij} = p_{i+} \cdot p_{+j}$ *para* $i = 1, \ldots, I$ *y para* $j = 1, \ldots, J.$

También, con la finalidad de mostrar que los parámetros de asociación son desviaciones del supuesto de independencia, muestran que están relacionados con los cocientes de ventajas para el caso 2×2 y el caso $2 \times J$ tal como veremos a continuación.

4.1 El caso 2×2

Agresti menciona en la sección "Interpretación de los parámetros" lo siguiente[4]:

"Existe una relación directa entre los cocientes de ventajas y los parámetros de asociación de los modelos logarítmico-lineales. La relación más sencilla se da en las tablas 2×2", la cual mostraremos a continuación.

Nos serán de utilidad las dos siguientes igualdades

$$\sum_{i=1}^{I} U_{12}(i,j) = 0 \text{ para } j = 1, \ldots, J$$

$$\sum_{j=1}^{J} U_{12}(i,j) = 0 \text{ para } i = 1, \ldots, I$$

que demostraremos a continuación:

$$
\begin{aligned}
\sum_{i=1}^{I} U_{12}(i,j) &= \sum_{i=1}^{I} \left(I_{ij} - \left(U + U_1(i) + U_2(j)\right)\right) \\
&= \sum_{i=1}^{I} I_{ij} - \sum_{i=1}^{I} U - \sum_{i=1}^{I} U_1(i) - \sum_{i=1}^{I} U_2(j) \\
&= \sum_{i=1}^{I} I_{ij} - \sum_{i=1}^{I} \frac{I_{++}}{IJ} - \sum_{i=1}^{I} \left(\frac{I_{i+}}{J} - \frac{I_{++}}{IJ}\right) - \sum_{i=1}^{I} \left(\frac{I_{+j}}{I} - \frac{I_{++}}{IJ}\right) \\
&= \sum_{i=1}^{I} I_{ij} - \sum_{i=1}^{I} \frac{I_{++}}{IJ} - \sum_{i=1}^{I} \frac{I_{i+}}{J} + \sum_{i=1}^{I} \frac{I_{++}}{IJ} - \sum_{i=1}^{I} \frac{I_{+j}}{I} + \sum_{i=1}^{I} \frac{I_{++}}{IJ} \\
&= \sum_{i=1}^{I} I_{ij} - \sum_{i=1}^{I} \frac{I_{++}}{IJ} - \sum_{i=1}^{I} \frac{\sum_{j=1}^{J} I_{ij}}{J} + \sum_{i=1}^{I} \frac{I_{++}}{IJ} - \sum_{i=1}^{I} \frac{I_{+j}}{I} + \sum_{i=1}^{I} \frac{I_{++}}{IJ} \\
&= \sum_{i=1}^{I} I_{ij} - \sum_{i=1}^{I} \frac{I_{++}}{IJ} - \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} I_{ij}}{J} + \sum_{i=1}^{I} \frac{I_{++}}{IJ} - \sum_{i=1}^{I} \frac{I_{+j}}{I} + \sum_{i=1}^{I} \frac{I_{++}}{IJ} \\
&= I_{+j} - I \cdot \frac{I_{++}}{IJ} - \frac{I_{++}}{J} + I \cdot \frac{I_{++}}{IJ} - I \cdot \frac{I_{+j}}{I} + I \cdot \frac{I_{++}}{IJ} \\
&= I_{+j} - \frac{I_{++}}{J} - \frac{I_{++}}{J} + \frac{I_{++}}{J} - I_{+j} + \frac{I_{++}}{J} = 0
\end{aligned}
$$

$$
\begin{aligned}
\sum_{j=1}^{J} U_{12}(i,j) &= \sum_{j=1}^{J} \left(I_{ij} - \left(U + U_1(i) + U_2(j)\right)\right) \\
&= \sum_{j=1}^{J} I_{ij} - \sum_{j=1}^{J} U - \sum_{j=1}^{J} U_1(i) - \sum_{j=1}^{J} U_2(j) \\
&= \sum_{j=1}^{J} I_{ij} - \sum_{j=1}^{J} \frac{I_{++}}{IJ} - \sum_{j=1}^{J} \left(\frac{I_{i+}}{J} - \frac{I_{++}}{IJ}\right) - \sum_{j=1}^{J} \left(\frac{I_{+j}}{I} - \frac{I_{++}}{IJ}\right) \\
&= \sum_{j=1}^{J} I_{ij} - \sum_{j=1}^{J} \frac{I_{++}}{IJ} - \sum_{j=1}^{J} \frac{I_{i+}}{J} + \sum_{j=1}^{J} \frac{I_{++}}{IJ} - \sum_{j=1}^{J} \frac{I_{+j}}{I} + \sum_{j=1}^{J} \frac{I_{++}}{IJ} \\
&= \sum_{j=1}^{J} I_{ij} - \sum_{j=1}^{J} \frac{I_{++}}{IJ} - \sum_{j=1}^{J} \frac{I_{i+}}{J} + \sum_{j=1}^{J} \frac{I_{++}}{IJ} - \sum_{j=1}^{J} \frac{\sum_{i=1}^{I} I_{ij}}{I} + \sum_{j=1}^{J} \frac{I_{++}}{IJ} \\
&= \sum_{j=1}^{J} I_{ij} - \sum_{j=1}^{J} \frac{I_{++}}{IJ} - \sum_{j=1}^{J} \frac{I_{i+}}{J} + \sum_{j=1}^{J} \frac{I_{++}}{IJ} - \frac{\sum_{j=1}^{J}\sum_{i=1}^{I} I_{ij}}{I} + \sum_{j=1}^{J} \frac{I_{++}}{IJ} \\
&= I_{i+} - J \cdot \frac{I_{++}}{IJ} - J \cdot \frac{I_{i+}}{J} + J \cdot \frac{I_{++}}{IJ} - \frac{I_{++}}{I} + J \cdot \frac{I_{++}}{IJ} \\
&= I_{i+} - \frac{I_{++}}{I} - I_{i+} + \frac{I_{++}}{I} - \frac{I_{++}}{I} + \frac{I_{++}}{I} = 0
\end{aligned}
$$

q.e.d.

En las tablas 2×2 tenemos que el cociente de ventajas es:

$$\theta = \frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}},$$

si le aplicamos la función logaritmo natural tenemos que

$$\begin{aligned}
\ln(\theta) = \ln\left(\frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}}\right) &= \ln(p_{11}) + \ln(p_{22}) - \ln(p_{21}) - \ln(p_{12}) \\
&= (U + U_1(1) + U_2(1) + U_{12}(1,1)) + (U + U_1(2) + U_2(2) + U_{12}(2,2)) \\
&\quad - (U + U_1(2) + U_2(1) + U_{12}(2,1)) - (U + U_1(1) + U_2(2) + U_{12}(1,2)) \\
&= U_{12}(1,1) + U_{12}(2,2) - U_{12}(2,1) - U_{12}(1,2)
\end{aligned}$$

pero, según lo que acabamos de demostrar, tenemos que

$$\begin{aligned}
\sum_{i=1}^{2} U_{12}(i,1) &= U_{12}(1,1) + U_{12}(2,1) = 0 \\
\sum_{i=1}^{2} U_{12}(i,2) &= U_{12}(1,2) + U_{12}(2,2) = 0 \\
\sum_{j=1}^{2} U_{12}(1,j) &= U_{12}(1,1) + U_{12}(1,2) = 0 \\
\sum_{j=1}^{2} U_{12}(2,j) &= U_{12}(2,1) + U_{12}(2,2) = 0
\end{aligned}$$

lo cual nos lleva a que

$$\begin{aligned}
\ln(\theta) &= U_{12}(1,1) + U_{12}(2,2) - U_{12}(2,1) - U_{12}(1,2) \\
&= U_{12}(1,1) + (-U_{12}(2,1)) - (-U_{12}(1,1)) - (-U_{12}(1,1)) \\
&= 3 \cdot U_{12}(1,1) - U_{12}(2,1) = 3 \cdot U_{12}(1,1) - (-U_{12}(1,1)) \\
&= 4 \cdot U_{12}(1,1) = -4 \cdot U_{12}(2,1) = -4 \cdot U_{12}(1,2) = 4 \cdot U_{12}(2,2)
\end{aligned}$$

que podemos resumir así:

$$U_{12}(i,j) = (-1)^{i+j} \frac{\ln(\theta)}{4}.$$

Esta relación muestra que en efecto los parámetros de asociación son desviaciones del supuesto de independencia pues

$$U_{12}(i,j) = 0 \ \ para \ \ i = 1,2 \ \ y \ \ para \ \ j = 1,2 \Leftrightarrow \theta = 1$$

$$\Leftrightarrow$$

$$p_{ij} = p_{i+} \cdot p_{+j} \ \ para \ \ i = 1,2 \ \ y \ \ para \ \ j = 1,2.$$

A continuación demostraremos está aseveración.

Demostraremos primero que

$$U_{12}(i,j) = 0 \ \ para \ \ i = 1,2 \ \ y \ \ para \ \ j = 1,2 \Leftrightarrow \theta = 1.$$

Por un lado

$$\theta = 1 \Rightarrow U_{12}(i,j) = (-1)^{i+j} \frac{\ln(\theta)}{4}$$

$$= (-1)^{i+j} \frac{\ln(1)}{4} = (-1)^{i+j} \frac{0}{4} = 0 \ \ para \ \ i = 1,2 \ \ y \ \ para \ \ j = 1,2,$$

por el otro

$$U_{12}(i,j) = 0 \ \ para \ \ i = 1,2 \ \ y \ \ para \ \ j = 1,2$$

$$\Rightarrow (-1)^{i+j} \frac{\ln(\theta)}{4} = 0 \Rightarrow \ln(\theta) = 0 \Rightarrow \theta = 1^*$$

q.e.d.

Ahora pasaremos a demostrar que

$$\theta = 1 \Leftrightarrow p_{ij} = p_{i+} \cdot p_{+j} \ \ para \ \ i = 1,2 \ \ y \ \ para \ \ j = 1,2.$$

Por un lado

$$p_{ij} = p_{i+} \cdot p_{+j} \ \ para \ \ i = 1,2 \ \ y \ \ para \ \ j = 1,2$$

$$\Rightarrow \theta = \frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}} = \frac{(p_{1+} \cdot p_{+1}) \cdot (p_{2+} \cdot p_{+2})}{(p_{2+} \cdot p_{+1}) \cdot (p_{1+} \cdot p_{+2})} = 1,$$

por otro lado

*Esta implicación se debe a que $\ln(\bullet)$ es una función biyectiva, en particular suprayectiva, y entonces tenemos que $\ln(\theta) = 0 \Rightarrow \theta = 1$ pues $\ln(1) = 0$.

$$\theta = \frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}} = 1 \Rightarrow p_{11} \cdot p_{22} = p_{21} \cdot p_{12}$$

$$\Rightarrow p_{11} \cdot p_{11} + p_{11} \cdot p_{12} + p_{11} \cdot p_{21} + p_{11} \cdot p_{22} = p_{11} \cdot p_{11} + p_{11} \cdot p_{12} + p_{11} \cdot p_{21} + p_{21} \cdot p_{12}$$

$$\Rightarrow p_{11} \cdot (p_{11} + p_{12} + p_{21} + p_{22}) = (p_{11} + p_{12}) \cdot (p_{11} + p_{21})$$

$$\Rightarrow p_{11} \cdot 1 = (p_{11} + p_{12}) \cdot (p_{11} + p_{21})$$

$$\Rightarrow p_{11} = p_{1+} \cdot p_{+1}$$

$$\theta = \frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}} = 1 \Rightarrow p_{11} \cdot p_{22} = p_{21} \cdot p_{12}$$

$$\Rightarrow p_{12} \cdot p_{11} + p_{12} \cdot p_{12} + p_{12} \cdot p_{21} + p_{12} \cdot p_{22} = p_{12} \cdot p_{11} + p_{12} \cdot p_{12} + p_{11} \cdot p_{22} + p_{12} \cdot p_{22}$$

$$\Rightarrow p_{12} \cdot (p_{11} + p_{12} + p_{21} + p_{22}) = (p_{11} + p_{12}) \cdot (p_{12} + p_{22})$$

$$\Rightarrow p_{12} \cdot 1 = (p_{11} + p_{12}) \cdot (p_{12} + p_{22})$$

$$\Rightarrow p_{12} = p_{1+} \cdot p_{+2}$$

$$\theta = \frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}} = 1 \Rightarrow p_{11} \cdot p_{22} = p_{21} \cdot p_{12}$$

$$\Rightarrow p_{21} \cdot p_{11} + p_{21} \cdot p_{12} + p_{21} \cdot p_{21} + p_{21} \cdot p_{22} = p_{21} \cdot p_{11} + p_{11} \cdot p_{22} + p_{21} \cdot p_{21} + p_{21} \cdot p_{22}$$

$$\Rightarrow p_{21} \cdot (p_{11} + p_{12} + p_{21} + p_{22}) = (p_{21} + p_{22}) \cdot (p_{11} + p_{21})$$

$$\Rightarrow p_{21} \cdot 1 = (p_{21} + p_{22}) \cdot (p_{11} + p_{21})$$

$$\Rightarrow p_{21} = p_{2+} \cdot p_{+1}$$

$$\theta = \frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}} = 1 \Rightarrow p_{11} \cdot p_{22} = p_{21} \cdot p_{12}$$

$$\Rightarrow p_{22} \cdot p_{11} + p_{22} \cdot p_{12} + p_{22} \cdot p_{21} + p_{22} \cdot p_{22} = p_{21} \cdot p_{12} + p_{22} \cdot p_{12} + p_{22} \cdot p_{21} + p_{22} \cdot p_{22}$$

$$\Rightarrow p_{22} \cdot (p_{11} + p_{12} + p_{21} + p_{22}) = (p_{21} + p_{22}) \cdot (p_{12} + p_{22})$$

$$\Rightarrow p_{22} \cdot 1 = (p_{21} + p_{22}) \cdot (p_{12} + p_{22})$$

$$\Rightarrow p_{22} = p_{2+} \cdot p_{+2}$$

q.e.d.

4.2 El caso $2 \times J$

Según Bishop et. al.[5] "Cuando el número de categorías por variable es mayor que 2, aún podemos expresar cada parámetro del modelo como función de los cocientes de ventajas. Primero consideremos el incremento del número de categorías de una variable únicamente." y presentan la siguiente igualdad:

$$\exp\left(U_{12}\left(1,1\right)\right) = \prod_{j=2}^{J}\left(\frac{m_{11}\cdot m_{2j}}{m_{21}\cdot m_{1j}}\right)^{\frac{1}{2J}}.$$

Aplicando la función logaritmo natural a ambos miembros de la igualdad y utilizando la notación E_{ij} que hemos utilizado en este escrito para el valor esperado en la tabla de contingencia podemos reescribir esta igualdad de la siguiente manera:

$$U_{12}\left(1,1\right) = \ln\left(\prod_{j=2}^{J}\left(\frac{E_{11}\cdot E_{2j}}{E_{21}\cdot E_{1j}}\right)^{\frac{1}{2J}}\right) = \frac{1}{2J}\ln\left(\prod_{j=2}^{J}\left(\frac{E_{11}\cdot E_{2j}}{E_{21}\cdot E_{1j}}\right)\right) = \frac{1}{2J}\sum_{j=2}^{J}\ln\left(\frac{E_{11}\cdot E_{2j}}{E_{21}\cdot E_{1j}}\right)$$

Más adelante veremos la generalización de esta relación para todos los parámetros de asociación $U_{12}(i,j)$ en el caso $2 \times J$.

4.3 El caso $I \times J$

La relación existente entre los parámetros de asociación del modelo logarítmico-lineal para el caso $I \times J$ y los cocientes de ventajas es la siguiente:

$$U_{12}\left(i,j\right) = \frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}\right)}{IJ}$$

donde

$$\theta_{ij,rs} = \frac{p_{ij}\cdot p_{rs}}{p_{rj}\cdot p_{is}}.$$

Veremos a continuación la demostración de esta igualdad.

Tenemos que

$$\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}\right)=\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\frac{p_{ij}\cdot p_{rs}}{p_{rj}\cdot p_{is}}\right)$$
$$=\sum_{r=1}^{I}\sum_{s=1}^{J}\left(\ln\left(p_{ij}\right)+\ln\left(p_{rs}\right)-\ln\left(p_{rj}\right)-\ln\left(p_{is}\right)\right)$$
$$=\sum_{r=1}^{I}\sum_{s=1}^{J}\left(I_{ij}+I_{rs}-I_{rj}-I_{is}\right)$$
$$=\sum_{r=1}^{I}\sum_{s=1}^{J}I_{ij}+\sum_{r=1}^{I}\sum_{s=1}^{J}I_{rs}-\sum_{r=1}^{I}\sum_{s=1}^{J}I_{rj}-\sum_{r=1}^{I}\sum_{s=1}^{J}I_{is}$$
$$=\sum_{r=1}^{I}\sum_{s=1}^{J}I_{ij}+\sum_{r=1}^{I}\sum_{s=1}^{J}I_{rs}-\sum_{s=1}^{J}I_{+j}-\sum_{r=1}^{I}I_{i+}$$
$$=IJ\cdot I_{ij}+I_{++}-J\cdot I_{+j}-I\cdot I_{i+}$$
$$=IJ\cdot\left(I_{ij}+\frac{I_{++}}{IJ}-\frac{I_{+j}}{I}-\frac{I_{i+}}{J}\right)$$
$$=IJ\cdot\left(I_{ij}-\frac{I_{++}}{IJ}-\frac{I_{i+}}{J}+\frac{I_{++}}{IJ}-\frac{I_{+j}}{I}+\frac{I_{++}}{IJ}\right)$$
$$=IJ\cdot\left(I_{ij}-\frac{I_{++}}{IJ}-\left(\frac{I_{i+}}{J}-\frac{I_{++}}{IJ}\right)-\left(\frac{I_{+j}}{I}-\frac{I_{++}}{IJ}\right)\right)$$
$$=IJ\cdot\left(I_{ij}-\left(\frac{I_{++}}{IJ}+\left(\frac{I_{+j}}{I}-\frac{I_{++}}{IJ}\right)+\left(\frac{I_{i+}}{J}-\frac{I_{++}}{IJ}\right)\right)\right)$$
$$=IJ\cdot\left(I_{ij}-\left(U+U_1\left(i\right)+U_2\left(j\right)\right)\right)$$
$$=IJ\cdot U_{12}\left(i,j\right)\Rightarrow U_{12}\left(i,j\right)=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}\right)}{IJ}$$

q.e.d.

Esta relación muestra que en efecto los parámetros de asociación son desviaciones del supuesto de independencia pues

$$U_{12}\left(i,j\right)=0\ \ para\ \ i=1,\ldots,I\ \ y\ \ para\ \ j=1,\ldots,J$$
$$\Leftrightarrow$$
$$\theta_{ij,rs}=1\ \ para\ \ r=1,\ldots,I\ ,\ \ para\ \ s=1,\ldots,J\ ,\ \ para\ \ i=1,\ldots,I\ \ y\ \ para\ \ j=1,\ldots,J$$
$$\Leftrightarrow$$
$$p_{ij}=p_{i+}\cdot p_{+j}\ \ para\ \ i=1,\ldots,I\ \ y\ \ para\ \ j=1,\ldots,J.$$

A continuación demostraremos está aseveración.

Demostraremos primero que

$$U_{12}(i,j)=0 \ \ para \ \ i=1,\ldots,I \ \ y \ \ para \ \ j=1,\ldots,J$$

$$\Leftrightarrow$$

$\theta_{ij,rs}=1$ $para$ $r=1,\ldots,I$, $para$ $s=1,\ldots,J$, $para$ $i=1,\ldots,I$ y $para$ $j=1,\ldots,J$.

Por un lado

$$\theta_{ij,rs}=1 \ \ para \ \ r=1,\ldots,I \ \ y \ \ para \ \ s=1,\ldots,J$$

$$\Rightarrow U_{12}(i,j)=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln(\theta_{ij,rs})}{IJ}=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln(1)}{IJ}=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}0}{IJ}=0$$

entonces

$\theta_{ij,rs}=1$ $para$ $r=1,\ldots,I$, $para$ $s=1,\ldots,J$, $para$ $i=1,\ldots,I$ y $para$ $j=1,\ldots,J$

$$\Rightarrow U_{12}(i,j)=0 \ \ para \ \ i=1,\ldots,I \ \ y \ \ para \ \ j=1,\ldots,J,$$

por el otro

$$U_{12}(i,j)=0 \ \ para \ \ i=1,\ldots,I \ \ y \ \ para \ \ j=1,\ldots,J$$

$$\Rightarrow I_{ij}=U+U_1(i)+U_2(j) \ \ para \ \ i=1,\ldots,I \ \ y \ \ para \ \ j=1,\ldots,J$$

$$\Rightarrow \ln(p_{ij})=U+U_1(i)+U_2(j) \ \ para \ \ i=1,\ldots,I \ \ y \ \ para \ \ j=1,\ldots,J$$

$$\Rightarrow p_{ij}=\exp(U+U_1(i)+U_2(j))$$

$$=\exp(U)\exp(U_1(i))\exp(U_2(j)) \ \ para \ \ i=1,\ldots,I \ \ y \ \ para \ \ j=1,\ldots,J \Rightarrow$$

$$\theta_{ij,rs}=\frac{p_{ij}\cdot p_{rs}}{p_{rj}\cdot p_{is}}=\frac{[\exp(U)\exp(U_1(i))\exp(U_2(j))]\cdot[\exp(U)\exp(U_1(r))\exp(U_2(s))]}{[\exp(U)\exp(U_1(r))\exp(U_2(j))]\cdot[\exp(U)\exp(U_1(i))\exp(U_2(s))]}=1$$

$para$ $r=1,\ldots,I$, $para$ $s=1,\ldots,J$, $para$ $i=1,\ldots,I$ y $para$ $j=1,\ldots,J$

q.e.d.

Ahora pasaremos a demostrar que

$\theta_{ij,rs} = 1$ *para* $r = 1, \ldots, I$, *para* $s = 1, \ldots, J$, *para* $i = 1, \ldots, I$ *y para* $j = 1, \ldots, J$

$$\Leftrightarrow$$

$$p_{ij} = p_{i+} \cdot p_{+j} \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J.$$

Por un lado

$\theta_{ij,rs} = 1$ *para* $r = 1, \ldots, I$, *para* $s = 1, \ldots, J$, *para* $i = 1, \ldots, I$ *y para* $j = 1, \ldots, J$

$$\Rightarrow U_{12}(i,j) = 0 \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

$$\Rightarrow I_{ij} = U + U_1(i) + U_2(j) \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

$$\Rightarrow \ln(p_{ij}) = U + U_1(i) + U_2(j) \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

$$\Rightarrow p_{ij} = \exp(U + U_1(i) + U_2(j))$$

$$= \exp(U)\exp(U_1(i))\exp(U_2(j)) \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J \Rightarrow$$

$$p_{i+} = \sum_{i=1}^{I} p_{ij} = \sum_{i=1}^{I}(\exp(U)\exp(U_1(i))\exp(U_2(j))) = \exp(U)\exp(U_2(j))\sum_{i=1}^{I}\exp(U_1(i))$$

$$p_{+j} = \sum_{j=1}^{J} p_{ij} = \sum_{j=1}^{J}(\exp(U)\exp(U_1(i))\exp(U_2(j))) = \exp(U)\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))$$

$$p_{++} = \sum_{i=1}^{I}\sum_{j=1}^{J} p_{ij} = \sum_{i=1}^{I}\sum_{j=1}^{J}(\exp(U)\exp(U_1(i))\exp(U_2(j))) = \exp(U)\sum_{i=1}^{I}\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))$$

$$\Rightarrow p_{i+} \cdot p_{+j} = \frac{p_{i+} \cdot p_{+j}}{1} = \frac{p_{i+} \cdot p_{+j}}{p_{++}}$$

$$= \frac{\left[\exp(U)\exp(U_2(j))\sum_{i=1}^{I}\exp(U_1(i))\right] \cdot \left[\exp(U)\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))\right]}{\exp(U)\sum_{i=1}^{I}\exp(U_1(i))\sum_{j=1}^{J}\exp(U_2(j))}$$

$$= \exp(U)\exp(U_1(i))\exp(U_2(j)) = p_{ij},$$

por otro lado

$$p_{ij} = p_{i+} \cdot p_{+j} \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

$$\Rightarrow \theta_{ij,rs} = \frac{p_{ij} \cdot p_{rs}}{p_{rj} \cdot p_{is}} = \frac{(p_{i+} \cdot p_{+j}) \cdot (p_{r+} \cdot p_{+s})}{(p_{r+} \cdot p_{+j}) \cdot (p_{i+} \cdot p_{+s})} = 1$$

para $r = 1, \ldots, I$, *para* $s = 1, \ldots, J$, *para* $i = 1, \ldots, I$ *y para* $j = 1, \ldots, J$

q.e.d.

Ahora veremos que la relación que demostramos, a saber,

$$U_{12}(i,j) = \frac{\sum_{r=1}^{I}\sum_{s=1}^{J} \ln(\theta_{ij,rs})}{IJ}$$

es equivalente, cuando $I=2$, a la que presentan Bishop et. al para el caso $2\times J$.

La relación presentada por Bishop et. al, la cual ya mencionamos, es la siguiente:

$$\exp(U_{12}(1,1)) = \prod_{j=2}^{J}\left(\frac{m_{11}\cdot m_{2j}}{m_{21}\cdot m_{1j}}\right)^{\frac{1}{2J}}$$

y podemos reescribirla de la siguiente manera:

$$\exp(U_{12}(1,1)) = \prod_{s=2}^{J}\left(\frac{E_{11}\cdot E_{2s}}{E_{21}\cdot E_{1s}}\right)^{\frac{1}{2J}} = 1\cdot\left(\prod_{s=2}^{J}\left(\frac{E_{11}\cdot E_{2s}}{E_{21}\cdot E_{1s}}\right)^{\frac{1}{2J}}\right)$$

$$= \left(\frac{E_{11}\cdot E_{21}}{E_{21}\cdot E_{11}}\right)^{\frac{1}{2J}}\cdot\left(\prod_{s=2}^{J}\left(\frac{E_{11}\cdot E_{2s}}{E_{21}\cdot E_{1s}}\right)^{\frac{1}{2J}}\right) = \prod_{s=1}^{J}\left(\frac{E_{11}\cdot E_{2s}}{E_{21}\cdot E_{1s}}\right)^{\frac{1}{2J}}$$

$$= 1^{\frac{1}{2J}}\cdot\left(\prod_{s=2}^{J}\left(\frac{E_{11}\cdot E_{2s}}{E_{21}\cdot E_{1s}}\right)^{\frac{1}{2J}}\right) = 1\cdot\prod_{s=1}^{J}\left(\frac{E_{11}\cdot E_{2s}}{E_{21}\cdot E_{1s}}\right)^{\frac{1}{2J}}$$

$$= \prod_{s=1}^{J}\left(\frac{E_{11}\cdot E_{1s}}{E_{11}\cdot E_{1s}}\right)^{\frac{1}{2J}}\prod_{s=1}^{J}\left(\frac{E_{11}\cdot E_{2s}}{E_{21}\cdot E_{1s}}\right)^{\frac{1}{2J}} = \prod_{r=1}^{2}\prod_{s=1}^{J}\left(\frac{E_{11}\cdot E_{rs}}{E_{r1}\cdot E_{1s}}\right)^{\frac{1}{2J}}$$

y aplicando la función logaritmo natural a ambos miembros de esta igualdad tenemos que

$$U_{12}(1,1) = \ln\left(\prod_{r=1}^{2}\prod_{s=1}^{J}\left(\frac{E_{11}\cdot E_{rs}}{E_{r1}\cdot E_{1s}}\right)^{\frac{1}{2J}}\right)$$

$$= \sum_{r=1}^{2}\sum_{s=1}^{J}\ln\left(\left(\frac{E_{11}\cdot E_{rs}}{E_{r1}\cdot E_{1s}}\right)^{\frac{1}{2J}}\right) = \sum_{r=1}^{2}\sum_{s=1}^{J}\frac{1}{2J}\ln\left(\frac{E_{11}\cdot E_{rs}}{E_{r1}\cdot E_{1s}}\right)$$

$$= \frac{\sum_{r=1}^{2}\sum_{s=1}^{J}\ln\left(\frac{n\cdot p_{11}\cdot n\cdot p_{rs}}{n\cdot p_{r1}\cdot n\cdot p_{1s}}\right)}{2J} = \frac{\sum_{r=1}^{2}\sum_{s=1}^{J}\ln\left(\frac{p_{11}\cdot p_{rs}}{p_{r1}\cdot p_{1s}}\right)}{2J}.$$

A continuación demostraremos que en el caso $2 \times J$:

$$U_{12}(i,j) = \frac{\sum_{r=1}^{2}\sum_{s=1}^{J} \ln(\theta_{ij,rs})}{2J}.$$

Tenemos que

$$\begin{aligned}
\sum_{r=1}^{2}\sum_{s=1}^{J} \ln(\theta_{ij,rs}) &= \sum_{r=1}^{2}\sum_{s=1}^{J} \ln\left(\frac{p_{ij}\cdot p_{rs}}{p_{rj}\cdot p_{is}}\right) \\
&= \sum_{r=1}^{2}\sum_{s=1}^{J} \left(\ln(p_{ij}) + \ln(p_{rs}) - \ln(p_{rj}) - \ln(p_{is})\right) \\
&= \sum_{r=1}^{2}\sum_{s=1}^{J} \left(I_{ij} + I_{rs} - I_{rj} - I_{is}\right) \\
&= \sum_{r=1}^{2}\sum_{s=1}^{J} I_{ij} + \sum_{r=1}^{2}\sum_{s=1}^{J} I_{rs} - \sum_{r=1}^{2}\sum_{s=1}^{J} I_{rj} - \sum_{r=1}^{2}\sum_{s=1}^{J} I_{is} \\
&= \sum_{r=1}^{2}\sum_{s=1}^{J} I_{ij} + \sum_{r=1}^{2}\sum_{s=1}^{J} I_{rs} - \sum_{s=1}^{J} I_{+j} - \sum_{r=1}^{2} I_{i+} \\
&= 2J\cdot I_{ij} + I_{++} - J\cdot I_{+j} - 2\cdot I_{i+} \\
&= 2J\cdot\left(I_{ij} + \frac{I_{++}}{2J} - \frac{I_{+j}}{2} - \frac{I_{i+}}{J}\right) \\
&= 2J\cdot\left(I_{ij} - \frac{I_{++}}{2J} - \frac{I_{i+}}{J} + \frac{I_{++}}{2J} - \frac{I_{+j}}{2} + \frac{I_{++}}{2J}\right) \\
&= 2J\cdot\left(I_{ij} - \frac{I_{++}}{2J} - \left(\frac{I_{i+}}{J} - \frac{I_{++}}{2J}\right) - \left(\frac{I_{+j}}{2} - \frac{I_{++}}{2J}\right)\right) \\
&= 2J\cdot\left(I_{ij} - \left(\frac{I_{++}}{2J} + \left(\frac{I_{+j}}{2} - \frac{I_{++}}{2J}\right) + \left(\frac{I_{i+}}{J} - \frac{I_{++}}{2J}\right)\right)\right) \\
&= 2J\cdot\left(I_{ij} - \left(U + U_1(i) + U_2(j)\right)\right) \\
&= 2J\cdot U_{12}(i,j) \Rightarrow U_{12}(i,j) = \frac{\sum_{r=1}^{2}\sum_{s=1}^{J} \ln(\theta_{ij,rs})}{2J}
\end{aligned}$$

q.e.d.

4.4 El caso $I \times J \times K$

En el segundo capítulo definimos el modelo logarítmico-lineal para tres variables categóricas de la siguiente manera:

$$I_{ijk} = U + U_1(i) + U_2(j) + U_3(k) + U_{12}(i,j) + U_{13}(i,k) + U_{23}(j,k) + U_{123}(i,j,k)$$

para $i = 1, \ldots, I$ *y para* $j = 1, \ldots, J$ *y para* $k = 1, \ldots, K$ donde

$$i)\quad U = \frac{\sum_{k=1}^{K} V^{(k)}}{K}$$

$$ii)\quad U_1(i) = \frac{\sum_{k=1}^{K} V_1^{(k)}(i)}{K}$$

$$iii)\quad U_2(j) = \frac{\sum_{k=1}^{K} V_2^{(k)}(j)}{K}$$

$$iv)\quad U_{12}(i,j) = \frac{\sum_{k=1}^{K} V_{12}^{(k)}(i,j)}{K}$$

$$v)\quad U_3(k) = V^{(k)} - U$$

$$vi)\quad U_{13}(i,k) = V_1^{(k)}(i) - U_1(i)$$

$$vii)\quad U_{23}(j,k) = V_2^{(k)}(j) - U_2(j)$$

$$viii)\quad U_{123}(i,j,k) = V_{12}^{(k)}(i,j) - U_{12}(i,j)$$

donde $V^{(k)}$, $V_1^{(k)}(i)$, $V_2^{(k)}(j)$ *y* $V_{12}^{(k)}(i,j)$ son los parámetros del modelo logarítmico-lineal para dos variables categóricas que se cumple en la categoría k de la tercera variable, es decir

$$i)\quad V^{(k)} = \frac{I_{++k}}{IJ} \quad \text{donde } I_{++k} = \sum_{i=1}^{I}\sum_{j=1}^{J} I_{ijk}$$

$$ii)\quad V_1^{(k)}(i) = \frac{I_{i+k}}{J} - \frac{I_{++k}}{IJ} \quad \text{donde } I_{i+k} = \sum_{j=1}^{J} I_{ijk} \quad \text{para } i = 1, \ldots, I$$

$$iii)\quad V_2^{(k)}(j) = \frac{I_{+jk}}{I} - \frac{I_{++k}}{IJ} \quad \text{donde } I_{+jk} = \sum_{i=1}^{I} I_{ijk} \quad \text{para } j = 1, \ldots, J$$

$$iv)\quad V_{12}^{(k)}(i,j) = I_{ijk} - \left(V^{(k)} + V_1^{(k)}(i) + V_2^{(k)}(j)\right) \quad \text{para } i = 1, \ldots, I \text{ y para } j = 1, \ldots, J$$

donde $I_{ijk} = \ln(p_{ijk})$ *para* $i = 1, \ldots, I$, *para* $j = 1, \ldots, J$ *y para* $k = 1, \ldots, K$.

En el segundo capítulo comentamos que

$$V_{12}^{(k)}(i,j)=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJ}$$

lo cual demostraremos a continuación:

$$\begin{aligned}
&\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)=\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\frac{p_{ijk}\cdot p_{rsk}}{p_{rjk}\cdot p_{isk}}\right)\\
&=\sum_{r=1}^{I}\sum_{s=1}^{J}\left(\ln\left(p_{ijk}\right)+\ln\left(p_{rsk}\right)-\ln\left(p_{rjk}\right)-\ln\left(p_{isk}\right)\right)\\
&=\sum_{r=1}^{I}\sum_{s=1}^{J}\left(I_{ijk}+I_{rsk}-I_{rjk}-I_{isk}\right)\\
&=\sum_{r=1}^{I}\sum_{s=1}^{J}I_{ijk}+\sum_{r=1}^{I}\sum_{s=1}^{J}I_{rsk}-\sum_{r=1}^{I}\sum_{s=1}^{J}I_{rjk}-\sum_{r=1}^{I}\sum_{s=1}^{J}I_{isk}\\
&=\sum_{r=1}^{I}\sum_{s=1}^{J}I_{ijk}+\sum_{r=1}^{I}\sum_{s=1}^{J}I_{rsk}-\sum_{s=1}^{J}I_{+jk}-\sum_{r=1}^{I}I_{i+k}\\
&=IJ\cdot I_{ijk}+I_{++k}-J\cdot I_{+jk}-I\cdot I_{i+k}\\
&=IJ\cdot\left(I_{ijk}+\frac{I_{++k}}{IJ}-\frac{I_{+jk}}{I}-\frac{I_{i+k}}{J}\right)\\
&=IJ\cdot\left(I_{ijk}-\frac{I_{++k}}{IJ}-\frac{I_{i+k}}{J}+\frac{I_{++k}}{IJ}-\frac{I_{+jk}}{I}+\frac{I_{++k}}{IJ}\right)\\
&=IJ\cdot\left(I_{ijk}-\frac{I_{++k}}{IJ}-\left(\frac{I_{i+k}}{J}-\frac{I_{++k}}{IJ}\right)-\left(\frac{I_{+jk}}{I}-\frac{I_{++k}}{IJ}\right)\right)\\
&=IJ\cdot\left(I_{ij}-\left(\frac{I_{++k}}{IJ}+\left(\frac{I_{+jk}}{I}-\frac{I_{++k}}{IJ}\right)+\left(\frac{I_{i+k}}{J}-\frac{I_{++k}}{IJ}\right)\right)\right)\\
&=IJ\cdot\left(I_{ijk}-\left(V^{(k)}+V_{1}^{(k)}(i)+V_{2}^{(k)}(j)\right)\right)\\
&=IJ\cdot V_{12}^{(k)}(i,j)\Rightarrow V_{12}^{(k)}(i,j)=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJ}
\end{aligned}$$

q.e.d.

También en el segundo capítulo comentamos que

$$U_{12}(i,j) = \frac{\sum_{k=1}^{K}\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK}$$

$$U_{13}(i,k) = \frac{\sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}\ln\left(\theta_{ik,rt}^{(j)}\right)}{IJK}$$

$$U_{23}(j,k) = \frac{\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\ln\left(\theta_{jk,st}^{(i)}\right)}{IJK}$$

donde

$$\theta_{ij,rs}^{(k)} = \frac{p_{ijk}\cdot p_{rsk}}{p_{rjk}\cdot p_{isk}}$$

$$\theta_{ik,rt}^{(j)} = \frac{p_{ijk}\cdot p_{rjt}}{p_{rjk}\cdot p_{ijt}}$$

$$\theta_{jk,st}^{(i)} = \frac{p_{ijk}\cdot p_{ist}}{p_{isk}\cdot p_{ijt}}$$

lo cual demostraremos a continuación.

Pasemos primero a demostrar que

$$U_{12}(i,j) = \frac{\sum_{k=1}^{K}\sum_{r=1}^{I}\sum_{sj=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK}.$$

Tenemos que

$$\frac{\sum_{k=1}^{K}\sum_{i=1}^{I}\sum_{j=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK} = \frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJ}$$

$$= \frac{1}{K}\sum_{k=1}^{K}V_{12}^{(k)}(i,j) = \frac{\sum_{k=1}^{K}V_{12}^{(k)}(i,j)}{K} = U_{12}(i,j)$$

q.e.d.

Ahora demostremos que

$$U_{13}(i,k) = \frac{\sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}\ln\left(\theta_{ik,rt}^{(j)}\right)}{IJK}.$$

Tenemos que

$$\sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}\ln\left(\theta_{ik,rt}^{(j)}\right) = \sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}\ln\left(\frac{p_{ijk}\cdot p_{rjt}}{p_{rjk}\cdot p_{ijt}}\right)$$

$$= \sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}\left(\ln\left(p_{ijk}\right)+\ln\left(p_{rjt}\right)-\ln\left(p_{rjk}\right)-\ln\left(p_{ijt}\right)\right)$$

$$= \sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}\left(I_{ijk}+I_{rjt}-I_{rjk}-I_{ijt}\right)$$

$$= \sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}I_{ijk}+\sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}I_{rjt}-\sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}I_{rjk}-\sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}I_{ijt}$$

$$= \sum_{r=1}^{I}\sum_{t=1}^{K}I_{i+k}+I_{+++}-\sum_{t=1}^{K}I_{++k}-\sum_{r=1}^{I}I_{i++}$$

$$= I\cdot K\cdot I_{i+k}+I_{+++}-K\cdot I_{++k}-I\cdot I_{i++}$$

$$= IJK\left(\frac{I_{i+k}}{J}+\frac{I_{+++}}{IJK}-\frac{I_{++k}}{IJ}-\frac{I_{i++}}{JK}\right)$$

$$= IJK\left(\left(\frac{I_{i+k}}{J}-\frac{I_{++k}}{IJ}\right)-\left(\frac{I_{i++}}{JK}-\frac{I_{+++}}{IJK}\right)\right)$$

$$= IJK\left(\left(\frac{I_{i+k}}{J}-\frac{I_{++k}}{IJ}\right)-\frac{1}{K}\left(\frac{I_{i++}}{J}-\frac{I_{+++}}{IJ}\right)\right)$$

$$= IJK\left(\left(\frac{I_{i+k}}{J}-\frac{I_{++k}}{IJ}\right)-\frac{1}{K}\sum_{k=1}^{K}\left(\frac{I_{i+k}}{J}-\frac{I_{++k}}{IJ}\right)\right)$$

$$= IJK\left(V_1^{(k)}(i)-\frac{1}{K}\sum_{k=1}^{K}V_1^{(k)}(i)\right) = IJK\left(V_1^{(k)}(i)-U_1(i)\right)$$

$$= IJK\cdot U_{13}(i,k) \Rightarrow U_{13}(i,k) = \frac{\sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}\ln\left(\theta_{ik,rt}^{(j)}\right)}{IJK}$$

q.e.d.

Ahora demostremos que

$$U_{23}(j,k)=\frac{\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\ln\left(\theta_{jk,st}^{(i)}\right)}{IJK}.$$

Tenemos que

$$\begin{aligned}
&\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\ln\left(\theta_{jk,st}^{(i)}\right)=\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\ln\left(\frac{p_{ijk}\cdot p_{ist}}{p_{isk}\cdot p_{ijt}}\right)\\
&=\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\left(\ln\left(p_{ijk}\right)+\ln\left(p_{ist}\right)-\ln\left(p_{isk}\right)-\ln\left(p_{ijt}\right)\right)\\
&=\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\left(I_{ijk}+I_{ist}-I_{isk}-I_{ijt}\right)\\
&=\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}I_{ijk}+\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}I_{ist}-\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}I_{isk}-\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}I_{ijt}\\
&=\sum_{s=1}^{J}\sum_{t=1}^{K}I_{+jk}+I_{+++}-\sum_{t=1}^{K}I_{++k}-\sum_{s=1}^{J}I_{+j+}\\
&=J\cdot K\cdot I_{+jk}+I_{+++}-K\cdot I_{++k}-J\cdot I_{+j+}\\
&=IJK\left(\frac{I_{+jk}}{I}+\frac{I_{+++}}{IJK}-\frac{I_{++k}}{IJ}-\frac{I_{+j+}}{IK}\right)\\
&=IJK\left(\left(\frac{I_{+jk}}{I}-\frac{I_{++k}}{IJ}\right)-\left(\frac{I_{+j+}}{IK}-\frac{I_{+++}}{IJK}\right)\right)\\
&=IJK\left(\left(\frac{I_{+jk}}{I}-\frac{I_{++k}}{IJ}\right)-\frac{1}{K}\sum_{k=1}^{K}\left(\frac{I_{+jk}}{I}-\frac{I_{++k}}{IJ}\right)\right)\\
&=IJK\left(V_2^{(k)}(j)-\frac{1}{K}\sum_{k=1}^{K}V_2^{(k)}(j)\right)\\
&=IJK\left(V_2^{(k)}(j)-U_2(j)\right)\\
&=IJK\cdot U_{23}(j,k)\Rightarrow U_{23}(j,k)=\frac{\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\ln\left(\theta_{jk,st}^{(i)}\right)}{IJK}
\end{aligned}$$

q.e.d.

REFERENCIAS

[1]Agresti, Alan, Categorical Data Analysis, pág. 133

[2]Ibidem., pág. 133

[3]Bishop, Yvonne M.M. et. al., Discrete Multivariate Analysis, pág. 18

[4]Agresti, Alan, op. cit., pág. 133

[5]Bishop, Yvonne M.M. et. al., op. cit., pág. 26

CONCLUSIONES

Primera

Es posible relacionar el cociente de ventajas con los parámetros de asociación del modelo logarítmico-lineal para dos variables categóricas en el caso 2×2 pues

$$U_{12}(i,j) = (-1)^{i+j} \frac{\ln(\theta)}{4} \quad para \quad i = 1,2 \quad y \quad para \quad j = 1,2$$

donde

$$\theta = \frac{p_{11} \cdot p_{22}}{p_{21} \cdot p_{12}}.$$

Esta relación muestra que los parámetros de asociación son desviaciones del supuesto de independencia pues

$$U_{12}(i,j) = 0 \quad para \quad i = 1,2 \quad y \quad para \quad j = 1,2$$

$$\Leftrightarrow$$

$$\theta = 1$$

$$\Leftrightarrow$$

$$p_{ij} = p_{i+} \cdot p_{+j} \quad para \quad i = 1,2 \quad y \quad para \quad j = 1,2.$$

Esta última igualdad implica la independencia entre ambas variables categóricas.

Segunda

Es posible extender dicha relación al caso $I \times J$ pues

$$U_{12}(i,j) = \frac{\sum_{r=1}^{I} \sum_{s=1}^{J} \ln(\theta_{ij,rs})}{IJ} \quad para \quad i = 1,\ldots,I \quad y \quad para \quad j = 1,\ldots,J$$

donde

$$\theta_{ij,rs} = \frac{p_{ij} \cdot p_{rs}}{p_{rj} \cdot p_{is}}.$$

Esta relación muestra que los parámetros de asociación son desviaciones del supuesto de independencia pues

$$U_{12}(i,j) = 0 \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

$$\Leftrightarrow$$

$$\theta_{ij,rs} = 1 \ \ para \ \ r = 1, \ldots, I \ , \ para \ \ s = 1, \ldots, J \ , \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

$$\Leftrightarrow$$

$$p_{ij} = p_{i+} \cdot p_{+j} \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J.$$

Esta última igualdad implica la independencia entre ambas variables categóricas.

Tercera

Interpretar los parámetros de asociación del modelo logarítmico-lineal para dos variables categóricas como desviaciones del supuesto de independencia ya era posible pues

$$U_{12}(i,j) = I_{ij} - (U + U_1(i) + U_2(j)) \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J,$$

pero reescribirlos de la siguiente manera

$$U_{12}(i,j) = \frac{\sum_{r=1}^{I} \sum_{s=1}^{J} \ln(\theta_{ij,rs})}{IJ}$$

permite, según veremos a continuación, una demostración formal de la siguiente igualdad:

$$I_{ij} = U + U_1(i) + U_2(j) + U_{12}(i,j) \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J.$$

No es válido, matemáticamente hablando, que para demostrar la igualdad

$$I_{ij} = U + U_1(i) + U_2(j) + U_{12}(i,j) \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

sean definidos los parámetros de asociación de la siguiente manera:

$$U_{12}(i,j) = I_{ij} - (U + U_1(i) + U_2(j)) \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

pues esta definición da por hecho la igualdad a demostrar.

En cambio sí es válido desde el principio definir los parámetros de asociación de la siguiente manera:

$$U_{12}(i,j)=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln(\theta_{ij,rs})}{IJ} \quad para \ \ i=1,\ldots,I \ \ y \ \ para \ \ j=1,\ldots,J$$

y después demostrar que

$$I_{ij}=U+U_1(i)+U_2(j)+U_{12}(i,j) \quad para \ \ i=1,\ldots,I \ \ y \ \ para \ \ j=1,\ldots,J.$$

Cuarta

Es posible extender la relación entre los cocientes de ventajas y los parámetros de asociación del modelo logarítmico-lineal para el caso $I\times J\times K$, pues

$$U_{12}(i,j)=\frac{\sum_{k=1}^{K}\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK}$$

$$U_{13}(i,k)=\frac{\sum_{j=1}^{J}\sum_{r=1}^{I}\sum_{t=1}^{K}\ln\left(\theta_{ik,rt}^{(j)}\right)}{IJK}$$

$$U_{23}(j,k)=\frac{\sum_{i=1}^{I}\sum_{s=1}^{J}\sum_{t=1}^{K}\ln\left(\theta_{jk,st}^{(i)}\right)}{IJK}$$

$$U_{123}(i,j,k)=\frac{\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJ}-\frac{\sum_{k=1}^{K}\sum_{r=1}^{I}\sum_{s=1}^{J}\ln\left(\theta_{ij,rs}^{(k)}\right)}{IJK}$$

donde

$$\theta_{ij,rs}^{(k)}=\frac{p_{ijk}\cdot p_{rsk}}{p_{rjk}\cdot p_{isk}}$$
$$\theta_{ik,rt}^{(j)}=\frac{p_{ijk}\cdot p_{rjt}}{p_{rjk}\cdot p_{ijt}}$$
$$\theta_{jk,st}^{(i)}=\frac{p_{ijk}\cdot p_{ist}}{p_{isk}\cdot p_{ijt}}.$$

Esta relación permite, según vimos en el capítulo segundo, demostrar fácilmente que

$$I_{ijk} = U + U_1 + U_2 + U_3 \Leftrightarrow p_{ijk} = p_{i++} \cdot p_{+j+} \cdot p_{++k}$$

Esta última igualdad implica la independencia conjunta de las tres variables categóricas.

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13} \Leftrightarrow p_{ijk} = p_{i+k} \cdot p_{+j+}$$

Esta última igualdad implica la independencia conjunta de la segunda variable de las otras dos.

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13} + U_{23} \Leftrightarrow p_{ijk} = \frac{p_{i+k} \cdot p_{+jk}}{p_{++k}}$$

Esta última igualdad implica la independencia condicional de la primera y segunda variable.

BIBLIOGRAFÍA

AGRESTI, Alan. Categorical Data Analysis.
Wiley Interscience, New York, 1990

BISHOP, Yvonne M.M. et. al. Discrete Multivariate Analysis.
MIT Press, Cambridge, 1975

MOOD, Alexander M. et. al. Introduction to the theory of statistics, third edition. McGraw Hill, Singapore, 1974

SHESKIN, David J. Handbook of parametric and non parametric statistical procedures, 2nd edition. Chapman&Hall, Boca Raton, 2000

VILLA PALATO, María Soledad. Análisis estadístico de la población escolar de primer ingreso a la ENEP Acatlán a través de modelos log-lineales. Tesis licenciatura (Actuario), ENEP Acatlán, 1986

APÉNDICE A

La tabla de contingencia que obtuvo Radelet en 1981 es la siguiente:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* Sí	No
Blanca	*Blanca*	19	132
	Negra	0	9
Negra	*Blanca*	11	52
	Negra	6	97

Las tablas marginales nos serán de utilidad para obtener los EMV de las frecuencias esperadas en la tabla de contingencia correspondientes a los diversos modelos.

Raza del acusado	O_{i++}
Blanca	160
Negra	166

Raza de la víctima	O_{+j+}
Blanca	214
Negra	112

Pena de muerte	O_{+k+}
Sí	36
No	290

Raza del acusado O_{ij+}	*Raza de la víctima* Blanca	Negra
Blanca	151	9
Negra	63	103

Raza del acusado O_{i+k}	*Pena de muerte* Sí	No
Blanca	19	141
Negra	17	149

Raza de la víctima O_{+jk}	*Pena de muerte* Sí	No
Blanca	30	184
Negra	6	106

Correspondiente al modelo de independencia entre raza del acusado (1) y raza de la víctima (2), raza del acusado y pena de muerte (3) e independencia entre pena de muerte y raza de la víctima:

$$I_{ijk} = U + U_1 + U_2 + U_3$$

tenemos la siguiente tabla de EMV de las frecuencias esperadas en la tabla de contingencia:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* Sí	No
Blanca	*Blanca*	11.60	93.43
	Negra	6.07	48.90
Negra	*Blanca*	12.03	96.94
	Negra	6.30	50.73

debido a que

$$\tilde{E}_{111} = n \cdot \tilde{p}_{111} = n \cdot \frac{O_{1++} \cdot O_{+1+} \cdot O_{++1}}{n^3} = 326 \cdot \frac{160 \cdot 214 \cdot 36}{326^3} = 11.5985$$
$$\tilde{E}_{112} = n \cdot \tilde{p}_{112} = n \cdot \frac{O_{1++} \cdot O_{+1+} \cdot O_{++2}}{n^3} = 326 \cdot \frac{160 \cdot 214 \cdot 290}{326^3} = 93.4322$$
$$\tilde{E}_{121} = n \cdot \tilde{p}_{121} = n \cdot \frac{O_{1++} \cdot O_{+2+} \cdot O_{++1}}{n^3} = 326 \cdot \frac{160 \cdot 112 \cdot 36}{326^3} = 6.0702$$
$$\tilde{E}_{122} = n \cdot \tilde{p}_{122} = n \cdot \frac{O_{1++} \cdot O_{+2+} \cdot O_{++2}}{n^3} = 326 \cdot \frac{160 \cdot 112 \cdot 290}{326^3} = 48.8991$$
$$\tilde{E}_{211} = n \cdot \tilde{p}_{211} = n \cdot \frac{O_{2++} \cdot O_{+1+} \cdot O_{++1}}{n^3} = 326 \cdot \frac{166 \cdot 214 \cdot 36}{326^3} = 12.0334$$
$$\tilde{E}_{212} = n \cdot \tilde{p}_{212} = n \cdot \frac{O_{2++} \cdot O_{+1+} \cdot O_{++2}}{n^3} = 326 \cdot \frac{166 \cdot 214 \cdot 290}{326^3} = 96.9359$$
$$\tilde{E}_{221} = n \cdot \tilde{p}_{221} = n \cdot \frac{O_{2++} \cdot O_{+2+} \cdot O_{++1}}{n^3} = 326 \cdot \frac{166 \cdot 112 \cdot 36}{326^3} = 6.2979$$
$$\tilde{E}_{222} = n \cdot \tilde{p}_{222} = n \cdot \frac{O_{2++} \cdot O_{+2+} \cdot O_{++2}}{n^3} = 326 \cdot \frac{166 \cdot 112 \cdot 290}{326^3} = 50.7328.$$

Correspondiente al modelo de independencia entre pena de muerte y raza de la víctima además de la independencia entre pena de muerte y raza del acusado:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{12}$$

tenemos la siguiente tabla de EMV de las frecuencias esperadas en la tabla de contingencia:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* *Sí*	*No*
Blanca	*Blanca*	16.67	134.33
	Negra	0.99	8.01
Negra	*Blanca*	6.96	56.04
	Negra	11.37	91.63

debido a que

$$\tilde{E}_{111} = n \cdot \tilde{p}_{111} = n \cdot \frac{O_{++1} \cdot O_{11+}}{n^2} = 326 \cdot \frac{36 \cdot 151}{326^2} = 16.6748$$
$$\tilde{E}_{112} = n \cdot \tilde{p}_{112} = n \cdot \frac{O_{++2} \cdot O_{11+}}{n^2} = 326 \cdot \frac{290 \cdot 151}{326^2} = 134.3252$$
$$\tilde{E}_{121} = n \cdot \tilde{p}_{121} = n \cdot \frac{O_{++1} \cdot O_{12+}}{n^2} = 326 \cdot \frac{36 \cdot 9}{326^2} = 0.9939$$
$$\tilde{E}_{122} = n \cdot \tilde{p}_{122} = n \cdot \frac{O_{++2} \cdot O_{12+}}{n^2} = 326 \cdot \frac{290 \cdot 9}{326^2} = 8.0061$$

$$\tilde{E}_{211} = n \cdot \tilde{p}_{211} = n \cdot \frac{O_{++1} \cdot O_{21+}}{n^2} = 326 \cdot \frac{36 \cdot 63}{326^2} = 6.9571$$
$$\tilde{E}_{212} = n \cdot \tilde{p}_{212} = n \cdot \frac{O_{++2} \cdot O_{21+}}{n^2} = 326 \cdot \frac{290 \cdot 63}{326^2} = 56.0409$$
$$\tilde{E}_{221} = n \cdot \tilde{p}_{221} = n \cdot \frac{O_{++1} \cdot O_{22+}}{n^2} = 326 \cdot \frac{36 \cdot 103}{326^2} = 11.3742$$
$$\tilde{E}_{222} = n \cdot \tilde{p}_{222} = n \cdot \frac{O_{++2} \cdot O_{22+}}{n^2} = 326 \cdot \frac{290 \cdot 103}{326^2} = 91.6258.$$

Correspondiente al modelo de independencia entre pena de muerte y raza de la víctima además de la independencia entre raza del acusado y raza de la víctima:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13}$$

tenemos la siguiente tabla de EMV de las frecuencias esperadas en la tabla de contingencia:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* *Sí*	*No*
Blanca	*Blanca*	12.47	92.56
	Negra	6.53	48.44
Negra	*Blanca*	11.16	97.81
	Negra	5.84	51.19

debido a que

$$\tilde{E}_{111} = n \cdot \tilde{p}_{111} = n \cdot \frac{O_{+1+} \cdot O_{1+1}}{n^2} = 326 \cdot \frac{214 \cdot 19}{326^2} = 12.4724$$
$$\tilde{E}_{112} = n \cdot \tilde{p}_{112} = n \cdot \frac{O_{+1+} \cdot O_{1+2}}{n^2} = 326 \cdot \frac{214 \cdot 141}{326^2} = 92.5583$$
$$\tilde{E}_{121} = n \cdot \tilde{p}_{121} = n \cdot \frac{O_{+2+} \cdot O_{1+1}}{n^2} = 326 \cdot \frac{112 \cdot 19}{326^2} = 6.5276$$
$$\tilde{E}_{122} = n \cdot \tilde{p}_{122} = n \cdot \frac{O_{+2+} \cdot O_{1+2}}{n^2} = 326 \cdot \frac{112 \cdot 141}{326^2} = 48.4417$$
$$\tilde{E}_{211} = n \cdot \tilde{p}_{211} = n \cdot \frac{O_{+1+} \cdot O_{2+1}}{n^2} = 326 \cdot \frac{214 \cdot 17}{326^2} = 11.1595$$
$$\tilde{E}_{212} = n \cdot \tilde{p}_{212} = n \cdot \frac{O_{+1+} \cdot O_{2+2}}{n^2} = 326 \cdot \frac{214 \cdot 149}{326^2} = 97.8098$$
$$\tilde{E}_{221} = n \cdot \tilde{p}_{221} = n \cdot \frac{O_{+2+} \cdot O_{2+1}}{n^2} = 326 \cdot \frac{112 \cdot 17}{326^2} = 5.8405$$
$$\tilde{E}_{222} = n \cdot \tilde{p}_{222} = n \cdot \frac{O_{+2+} \cdot O_{2+2}}{n^2} = 326 \cdot \frac{112 \cdot 149}{326^2} = 51.1902.$$

Correspondiente al modelo de independencia entre pena de muerte y raza del acusado además de la independencia entre raza del acusado y raza de la víctima:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{23}$$

tenemos la siguiente tabla de EMV de las frecuencias esperadas en la tabla de contingencia:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* Sí	No
Blanca	*Blanca*	14.72	90.31
	Negra	2.94	52.02
Negra	*Blanca*	15.28	93.69
	Negra	3.06	53.98

debido a que

$$\tilde{E}_{111} = n \cdot \tilde{p}_{111} = n \cdot \frac{O_{1++} \cdot O_{+11}}{n^2} = 326 \cdot \frac{160 \cdot 30}{326^2} = 14.7239$$
$$\tilde{E}_{112} = n \cdot \tilde{p}_{112} = n \cdot \frac{O_{1++} \cdot O_{+12}}{n^2} = 326 \cdot \frac{160 \cdot 184}{326^2} = 90.3067$$
$$\tilde{E}_{121} = n \cdot \tilde{p}_{121} = n \cdot \frac{O_{1++} \cdot O_{+21}}{n^2} = 326 \cdot \frac{160 \cdot 6}{326^2} = 2.9448$$
$$\tilde{E}_{122} = n \cdot \tilde{p}_{122} = n \cdot \frac{O_{1++} \cdot O_{+22}}{n^2} = 326 \cdot \frac{160 \cdot 106}{326^2} = 52.0245$$
$$\tilde{E}_{211} = n \cdot \tilde{p}_{211} = n \cdot \frac{O_{2++} \cdot O_{+11}}{n^2} = 326 \cdot \frac{166 \cdot 30}{326^2} = 15.2761$$
$$\tilde{E}_{212} = n \cdot \tilde{p}_{212} = n \cdot \frac{O_{2++} \cdot O_{+12}}{n^2} = 326 \cdot \frac{166 \cdot 184}{326^2} = 93.6933$$
$$\tilde{E}_{221} = n \cdot \tilde{p}_{221} = n \cdot \frac{O_{2++} \cdot O_{+21}}{n^2} = 326 \cdot \frac{166 \cdot 6}{326^2} = 3.0552$$
$$\tilde{E}_{222} = n \cdot \tilde{p}_{222} = n \cdot \frac{O_{2++} \cdot O_{+22}}{n^2} = 326 \cdot \frac{166 \cdot 106}{326^2} = 53.9755.$$

Correspondiente al modelo de independencia condicional entre pena de muerte y raza de la víctima:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{12} + U_{13}$$

tenemos la siguiente tabla de EMV de las frecuencias esperadas en la tabla de contingencia:

Raza del acusado	Raza de la víctima	Pena de muerte Sí	No
Blanca	*Blanca*	17.93	133.07
	Negra	1.07	7.93
Negra	*Blanca*	6.45	56.55
	Negra	10.55	92.45

debido a que

$$\tilde{E}_{111} = n \cdot \tilde{p}_{111} = n \cdot \frac{O_{11+} \cdot O_{1+1}}{n \cdot O_{1++}} = 326 \cdot \frac{151 \cdot 19}{326 \cdot 160} = 17.9312$$
$$\tilde{E}_{112} = n \cdot \tilde{p}_{112} = n \cdot \frac{O_{11+} \cdot O_{1+2}}{n \cdot O_{1++}} = 326 \cdot \frac{151 \cdot 141}{326 \cdot 160} = 133.0687$$
$$\tilde{E}_{121} = n \cdot \tilde{p}_{121} = n \cdot \frac{O_{12+} \cdot O_{1+1}}{n \cdot O_{1++}} = 326 \cdot \frac{9 \cdot 19}{326 \cdot 160} = 1.0688$$
$$\tilde{E}_{122} = n \cdot \tilde{p}_{122} = n \cdot \frac{O_{12+} \cdot O_{1+2}}{n \cdot O_{1++}} = 326 \cdot \frac{9 \cdot 141}{326 \cdot 160} = 7.9313$$
$$\tilde{E}_{211} = n \cdot \tilde{p}_{211} = n \cdot \frac{O_{21+} \cdot O_{2+1}}{n \cdot O_{2++}} = 326 \cdot \frac{63 \cdot 17}{326 \cdot 166} = 6.4518$$
$$\tilde{E}_{212} = n \cdot \tilde{p}_{212} = n \cdot \frac{O_{21+} \cdot O_{2+2}}{n \cdot O_{2++}} = 326 \cdot \frac{63 \cdot 149}{326 \cdot 166} = 56.5482$$
$$\tilde{E}_{221} = n \cdot \tilde{p}_{221} = n \cdot \frac{O_{22+} \cdot O_{2+1}}{n \cdot O_{2++}} = 326 \cdot \frac{103 \cdot 17}{326 \cdot 166} = 10.5482$$
$$\tilde{E}_{222} = n \cdot \tilde{p}_{222} = n \cdot \frac{O_{22+} \cdot O_{2+2}}{n \cdot O_{2++}} = 326 \cdot \frac{103 \cdot 149}{326 \cdot 166} = 92.4518.$$

Correspondiente al modelo de independencia condicional entre pena de muerte y raza del acusado:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{12} + U_{23}$$

tenemos la siguiente tabla de EMV de las frecuencias esperadas en la tabla de contingencia:

Raza del acusado	Raza de la víctima	Pena de muerte Sí	No
Blanca	*Blanca*	21.17	129.83
	Negra	0.48	8.52
Negra	*Blanca*	8.83	54.17
	Negra	5.52	97.48

debido a que

$$\tilde{E}_{111} = n \cdot \tilde{p}_{111} = n \cdot \frac{O_{11+} \cdot O_{+11}}{n \cdot O_{+1+}} = 326 \cdot \frac{151 \cdot 30}{326 \cdot 214} = 21.1682$$
$$\tilde{E}_{112} = n \cdot \tilde{p}_{112} = n \cdot \frac{O_{11+} \cdot O_{+12}}{n \cdot O_{+1+}} = 326 \cdot \frac{151 \cdot 184}{326 \cdot 214} = 129.8318$$
$$\tilde{E}_{121} = n \cdot \tilde{p}_{121} = n \cdot \frac{O_{12+} \cdot O_{+21}}{n \cdot O_{+2+}} = 326 \cdot \frac{9 \cdot 6}{326 \cdot 112} = 0.4821$$
$$\tilde{E}_{122} = n \cdot \tilde{p}_{122} = n \cdot \frac{O_{12+} \cdot O_{+22}}{n \cdot O_{+2+}} = 326 \cdot \frac{9 \cdot 106}{326 \cdot 112} = 8.5179$$
$$\tilde{E}_{211} = n \cdot \tilde{p}_{211} = n \cdot \frac{O_{21+} \cdot O_{+11}}{n \cdot O_{+1+}} = 326 \cdot \frac{63 \cdot 30}{326 \cdot 214} = 8.8318$$
$$\tilde{E}_{212} = n \cdot \tilde{p}_{212} = n \cdot \frac{O_{21+} \cdot O_{+12}}{n \cdot O_{+1+}} = 326 \cdot \frac{63 \cdot 184}{326 \cdot 214} = 54.1682$$
$$\tilde{E}_{221} = n \cdot \tilde{p}_{221} = n \cdot \frac{O_{22+} \cdot O_{+21}}{n \cdot O_{+2+}} = 326 \cdot \frac{103 \cdot 6}{326 \cdot 112} = 5.5179$$
$$\tilde{E}_{222} = n \cdot \tilde{p}_{222} = n \cdot \frac{O_{22+} \cdot O_{+22}}{n \cdot O_{+2+}} = 326 \cdot \frac{103 \cdot 106}{326 \cdot 112} = 97.4821.$$

Correspondiente al modelo de independencia condicional entre raza del acusado y raza de la víctima:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{13} + U_{23}$$

tenemos la siguiente tabla de EMV de las frecuencias esperadas en la tabla de contingencia:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* *Sí*	*No*
Blanca	*Blanca*	15.83	89.46
	Negra	3.17	51.54
Negra	*Blanca*	14.17	94.54
	Negra	2.83	54.46

debido a que

$$\tilde{E}_{111} = n \cdot \tilde{p}_{111} = n \cdot \frac{O_{1+1} \cdot O_{+11}}{n \cdot O_{++1}} = 326 \cdot \frac{19 \cdot 30}{326 \cdot 36} = 15.8333$$
$$\tilde{E}_{112} = n \cdot \tilde{p}_{112} = n \cdot \frac{O_{1+2} \cdot O_{+12}}{n \cdot O_{++2}} = 326 \cdot \frac{141 \cdot 184}{326 \cdot 290} = 89.4621$$
$$\tilde{E}_{121} = n \cdot \tilde{p}_{121} = n \cdot \frac{O_{1+1} \cdot O_{+21}}{n \cdot O_{++1}} = 326 \cdot \frac{19 \cdot 6}{326 \cdot 36} = 3.1667$$
$$\tilde{E}_{122} = n \cdot \tilde{p}_{122} = n \cdot \frac{O_{1+2} \cdot O_{+22}}{n \cdot O_{++2}} = 326 \cdot \frac{141 \cdot 106}{326 \cdot 290} = 51.5379$$

$$\tilde{E}_{211} = n \cdot \tilde{p}_{211} = n \cdot \frac{O_{2+1} \cdot O_{+11}}{n \cdot O_{++1}} = 326 \cdot \frac{17 \cdot 30}{326 \cdot 36} = 14.1667$$
$$\tilde{E}_{212} = n \cdot \tilde{p}_{212} = n \cdot \frac{O_{2+2} \cdot O_{+12}}{n \cdot O_{++2}} = 326 \cdot \frac{149 \cdot 184}{326 \cdot 290} = 94.5379$$
$$\tilde{E}_{221} = n \cdot \tilde{p}_{221} = n \cdot \frac{O_{2+1} \cdot O_{+21}}{n \cdot O_{++1}} = 326 \cdot \frac{17 \cdot 6}{326 \cdot 36} = 2.8333$$
$$\tilde{E}_{222} = n \cdot \tilde{p}_{222} = n \cdot \frac{O_{2+2} \cdot O_{+22}}{n \cdot O_{++2}} = 326 \cdot \frac{149 \cdot 106}{326 \cdot 290} = 54.4621.$$

Correspondiente al modelo de dependencia condicional entre raza del acusado y raza de la víctima, raza del acusado y pena de muerte y raza de la víctima y pena de muerte:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{12} + U_{13} + U_{23}$$

tenemos la siguiente tabla de EMV de las frecuencias esperadas en la tabla de contingencia:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* *Sí*	*No*
Blanca	*Blanca*	18.67	132.33
	Negra	0.33	8.67
Negra	*Blanca*	11.33	51.67
	Negra	5.67	97.33

la cual se obtiene al culminar el octavo ciclo del ajuste proporcional iterativo con $\delta = 0.0001$. A continuación detallamos los cálculos del primer ciclo:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* *Sí*	*No*
Blanca	*Blanca*	$\tilde{E}_{111}^{(0)} = 1$	$\tilde{E}_{112}^{(0)} = 1$
	Negra	$\tilde{E}_{121}^{(0)} = 1$	$\tilde{E}_{122}^{(0)} = 1$
Negra	*Blanca*	$\tilde{E}_{211}^{(0)} = 1$	$\tilde{E}_{212}^{(0)} = 1$
	Negra	$\tilde{E}_{221}^{(0)} = 1$	$\tilde{E}_{222}^{(0)} = 1$

Raza del acusado	*Raza de la víctima* *Blanca*	*Negra*
Blanca	$\tilde{E}_{11+}^{(0)} = \tilde{E}_{111}^{(0)} + \tilde{E}_{112}^{(0)} = 2$	$\tilde{E}_{12+}^{(0)} = \tilde{E}_{121}^{(0)} + \tilde{E}_{122}^{(0)} = 2$
Negra	$\tilde{E}_{21+}^{(0)} = \tilde{E}_{211}^{(0)} + \tilde{E}_{212}^{(0)} = 2$	$\tilde{E}_{22+}^{(0)} = \tilde{E}_{221}^{(0)} + \tilde{E}_{222}^{(0)} = 2$

D	V	P: Sí	P: No
B	B	$\tilde{E}^{(1)}_{111} = \tilde{E}^{(0)}_{111} \cdot \frac{O_{11+}}{\tilde{E}^{(0)}_{11+}} = 1 \cdot \frac{151}{2} = 75.5$	$\tilde{E}^{(1)}_{112} = \tilde{E}^{(0)}_{112} \cdot \frac{O_{11+}}{\tilde{E}^{(0)}_{11+}} = 1 \cdot \frac{151}{2} = 75.5$
	N	$\tilde{E}^{(1)}_{121} = \tilde{E}^{(0)}_{121} \cdot \frac{O_{12+}}{\tilde{E}^{(0)}_{12+}} = 1 \cdot \frac{9}{2} = 4.5$	$\tilde{E}^{(1)}_{122} = \tilde{E}^{(0)}_{122} \cdot \frac{O_{12+}}{\tilde{E}^{(0)}_{12+}} = 1 \cdot \frac{9}{2} = 4.5$
N	B	$\tilde{E}^{(1)}_{211} = \tilde{E}^{(0)}_{211} \cdot \frac{O_{21+}}{\tilde{E}^{(0)}_{21+}} = 1 \cdot \frac{63}{2} = 31.5$	$\tilde{E}^{(1)}_{212} = \tilde{E}^{(0)}_{212} \cdot \frac{O_{21+}}{\tilde{E}^{(0)}_{21+}} = 1 \cdot \frac{63}{2} = 31.5$
	N	$\tilde{E}^{(1)}_{221} = \tilde{E}^{(0)}_{221} \cdot \frac{O_{22+}}{\tilde{E}^{(0)}_{22+}} = 1 \cdot \frac{103}{2} = 51.5$	$\tilde{E}^{(1)}_{222} = \tilde{E}^{(0)}_{222} \cdot \frac{O_{22+}}{\tilde{E}^{(0)}_{22+}} = 1 \cdot \frac{103}{2} = 51.5$

Raza del acusado	*Pena de muerte*: *Sí*	*No*
Blanca	$\tilde{E}^{(1)}_{1+1} = \tilde{E}^{(1)}_{111} + \tilde{E}^{(1)}_{121} = 80$	$\tilde{E}^{(1)}_{1+2} = \tilde{E}^{(1)}_{112} + \tilde{E}^{(1)}_{122} = 80$
Negra	$\tilde{E}^{(1)}_{2+1} = \tilde{E}^{(1)}_{211} + \tilde{E}^{(1)}_{221} = 83$	$\tilde{E}^{(1)}_{2+2} = \tilde{E}^{(1)}_{212} + \tilde{E}^{(1)}_{222} = 83$

D	V	P: Sí	P: No
B	B	$\tilde{E}^{(2)}_{111} = \tilde{E}^{(1)}_{111} \cdot \frac{O_{1+1}}{\tilde{E}^{(1)}_{1+1}} = 75.5 \cdot \frac{19}{80} = 17.9312$	$\tilde{E}^{(2)}_{112} = \tilde{E}^{(1)}_{112} \cdot \frac{O_{1+2}}{\tilde{E}^{(1)}_{1+2}} = 75.5 \cdot \frac{141}{80} = 133.0687$
	N	$\tilde{E}^{(2)}_{121} = \tilde{E}^{(1)}_{121} \cdot \frac{O_{1+1}}{\tilde{E}^{(1)}_{1+1}} = 4.5 \cdot \frac{19}{80} = 1.0687$	$\tilde{E}^{(2)}_{122} = \tilde{E}^{(1)}_{122} \cdot \frac{O_{1+2}}{\tilde{E}^{(1)}_{1+2}} = 4.5 \cdot \frac{141}{80} = 7.9312$
N	B	$\tilde{E}^{(2)}_{211} = \tilde{E}^{(1)}_{211} \cdot \frac{O_{2+1}}{\tilde{E}^{(1)}_{2+1}} = 31.5 \cdot \frac{17}{83} = 6.4518$	$\tilde{E}^{(2)}_{212} = \tilde{E}^{(1)}_{212} \cdot \frac{O_{2+2}}{\tilde{E}^{(1)}_{2+2}} = 31.5 \cdot \frac{149}{83} = 56.5482$
	N	$\tilde{E}^{(2)}_{221} = \tilde{E}^{(1)}_{221} \cdot \frac{O_{2+1}}{\tilde{E}^{(1)}_{2+1}} = 51.5 \cdot \frac{17}{83} = 10.5482$	$\tilde{E}^{(2)}_{222} = \tilde{E}^{(1)}_{222} \cdot \frac{O_{2+2}}{\tilde{E}^{(1)}_{2+2}} = 51.5 \cdot \frac{149}{83} = 92.4518$

Raza de la víctima	*Pena de muerte*: *Sí*	*No*
Blanca	$\tilde{E}^{(2)}_{+11} = \tilde{E}^{(2)}_{111} + \tilde{E}^{(2)}_{211} = 24.3830$	$\tilde{E}^{(2)}_{+12} = \tilde{E}^{(2)}_{112} + \tilde{E}^{(2)}_{212} = 189.6169$
Negra	$\tilde{E}^{(2)}_{+21} = \tilde{E}^{(2)}_{121} + \tilde{E}^{(2)}_{221} = 11.6169$	$\tilde{E}^{(2)}_{+22} = \tilde{E}^{(2)}_{122} + \tilde{E}^{(2)}_{222} = 100.3830$

$$\tilde{E}_{111}^{(3)} = \tilde{E}_{111}^{(2)} \cdot \frac{O_{+11}}{\tilde{E}_{+11}^{(2)}} = \frac{17.9312 \cdot 30}{24.3830} = 22.0619 \quad \tilde{E}_{112}^{(3)} = \tilde{E}_{112}^{(2)} \cdot \frac{O_{+12}}{\tilde{E}_{+12}^{(2)}} = \frac{133.0687 \cdot 184}{189.6169} = 129.1269$$

$$\tilde{E}_{121}^{(3)} = \tilde{E}_{121}^{(2)} \cdot \frac{O_{+21}}{\tilde{E}_{+21}^{(2)}} = \frac{1.0687 \cdot 6}{11.6169} = 0.5520 \quad \tilde{E}_{122}^{(3)} = \tilde{E}_{122}^{(2)} \cdot \frac{O_{+22}}{\tilde{E}_{+22}^{(2)}} = \frac{7.9312 \cdot 106}{100.3830} = 8.3750$$

$$\tilde{E}_{211}^{(3)} = \tilde{E}_{211}^{(2)} \cdot \frac{O_{+11}}{\tilde{E}_{+11}^{(2)}} = \frac{6.4518 \cdot 30}{24.3830} = 7.9381 \quad \tilde{E}_{212}^{(3)} = \tilde{E}_{212}^{(2)} \cdot \frac{O_{+12}}{\tilde{E}_{+12}^{(2)}} = \frac{56.5482 \cdot 184}{189.6169} = 54.8731$$

$$\tilde{E}_{221}^{(3)} = \tilde{E}_{221}^{(2)} \cdot \frac{O_{+21}}{\tilde{E}_{+21}^{(2)}} = \frac{10.5482 \cdot 6}{11.6169} = 5.4480 \quad \tilde{E}_{222}^{(3)} = \tilde{E}_{222}^{(2)} \cdot \frac{O_{+22}}{\tilde{E}_{+22}^{(2)}} = \frac{92.4518 \cdot 106}{100.3830} = 97.6250$$

El segundo ciclo comienza con los valores iniciales:

$$\tilde{E}_{ijk}^{(0)} = \tilde{E}_{ijk}^{(3)} \quad para\ \ i = 1,2, \ \ para\ \ j = 1,2 \ \ y\ \ para\ \ k = 1,2$$

y se lleva a cabo el procedimiento del primer ciclo, y así sucesivamente hasta completar ocho ciclos que es cuando se logra la precisión:

$$\left|\tilde{E}_{ijk}^{(3.8)} - \tilde{E}_{ijk}^{(3.7)}\right| < 0.0001 \quad para\ \ i = 1,2, \ \ para\ \ j = 1,2 \ \ y\ \ para\ \ k = 1,2.$$

Ya sólo resta por ejemplificar la estimación para el modelo saturado:

$$I_{ijk} = U + U_1 + U_2 + U_3 + U_{12} + U_{13} + U_{23} + U_{123}$$

al cual corresponde la tabla de contingencia:

Raza del acusado	*Raza de la víctima*	*Pena de muerte* *Sí*	*No*
Blanca	*Blanca*	19	132
	Negra	0	9
Negra	*Blanca*	11	52
	Negra	6	97

debido a que

$$\tilde{E}_{ijk} = n \cdot \tilde{p}_{ijk} = n \cdot \frac{O_{ijk}}{n} = O_{ijk} \quad para\ \ i = 1,2, \ \ para\ \ j = 1,2 \ \ y\ \ para\ \ k = 1,2.$$

APÉNDICE B

Para el modelo (1,2,3) tenemos que

$$U_{12}(i,j) = 0 \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ j = 1, \ldots, J$$

$$U_{13}(i,k) = 0 \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ k = 1, \ldots, K$$

$$U_{23}(j,k) = 0 \ \ para \ \ j = 1, \ldots, J \ \ y \ \ para \ \ k = 1, \ldots, K$$

$U_{123}(i,j,k) = 0$ *para* $i = 1, \ldots, I$, *para* $j = 1, \ldots, J$ *y para* $k = 1, \ldots, K$

y por lo tanto el número de parámetros nulos es

$$I \cdot J + I \cdot K + J \cdot K + I \cdot J \cdot K$$

pero, según mencionamos en el capítulo anterior,

$$\sum_{i=1}^{I} U_{12}(i,j) = \sum_{j=1}^{J} U_{12}(i,j) = 0$$
$$\sum_{i=1}^{I} U_{13}(i,k) = \sum_{k=1}^{K} U_{13}(i,k) = 0$$
$$\sum_{j=1}^{J} U_{23}(j,k) = \sum_{k=1}^{K} U_{23}(j,k) = 0$$
$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

y por lo tanto, en este modelo, los parámetros nulos independientes son

$$(I-1)\cdot(J-1)+(I-1)\cdot(K-1)+(J-1)\cdot(K-1)+(I-1)\cdot(J-1)\cdot(K-1).$$

Para el modelo (12,3) tenemos que

$$U_{13}(i,k) = 0 \ \ para \ \ i = 1, \ldots, I \ \ y \ \ para \ \ k = 1, \ldots, K$$

$$U_{23}(j,k) = 0 \ \ para \ \ j = 1, \ldots, J \ \ y \ \ para \ \ k = 1, \ldots, K$$

$U_{123}(i,j,k) = 0$ *para* $i = 1, \ldots, I$, *para* $j = 1, \ldots, J$ *y para* $k = 1, \ldots, K$

y por lo tanto el número de parámetros nulos es

$$I \cdot K + J \cdot K + I \cdot J \cdot K$$

pero, según mencionamos en el capítulo anterior,

$$\sum_{i=1}^{I} U_{13}(i,k) = \sum_{k=1}^{K} U_{13}(i,k) = 0$$
$$\sum_{j=1}^{J} U_{23}(j,k) = \sum_{k=1}^{K} U_{23}(j,k) = 0$$
$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

y por lo tanto, en este modelo, los parámetros nulos independientes son

$$(I-1)\cdot(K-1)+(J-1)\cdot(K-1)+(I-1)\cdot(J-1)\cdot(K-1).$$

Para el modelo $(13,2)$ tenemos que

$$U_{12}(i,j) = 0 \text{ } para \text{ } i = 1,\ldots,I \text{ } y \text{ } para \text{ } j = 1,\ldots,J$$

$$U_{23}(j,k) = 0 \text{ } para \text{ } j = 1,\ldots,J \text{ } y \text{ } para \text{ } k = 1,\ldots,K$$

$U_{123}(i,j,k) = 0$ *para* $i = 1,\ldots,I$, *para* $j = 1,\ldots,J$ *y para* $k = 1,\ldots,K$

y por lo tanto el número de parámetros nulos es

$$I\cdot J + J\cdot K + I\cdot J\cdot K$$

pero, según mencionamos en el capítulo anterior,

$$\sum_{i=1}^{I} U_{12}(i,j) = \sum_{j=1}^{J} U_{12}(i,j) = 0$$
$$\sum_{j=1}^{J} U_{23}(j,k) = \sum_{k=1}^{K} U_{23}(j,k) = 0$$
$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

y por lo tanto, en este modelo, los parámetros nulos independientes son

$$(I-1)\cdot(J-1)+(J-1)\cdot(K-1)+(I-1)\cdot(J-1)\cdot(K-1).$$

Para el modelo $(23, 1)$ tenemos que

$$U_{12}(i,j) = 0 \ para \ i = 1,\ldots,I \ y \ para \ j = 1,\ldots,J$$

$$U_{13}(i,k) = 0 \ para \ i = 1,\ldots,I \ y \ para \ k = 1,\ldots,K$$

$$U_{123}(i,j,k) = 0 \ para \ i = 1,\ldots,I \ , \ para \ j = 1,\ldots,J \ y \ para \ k = 1,\ldots,K$$

y por lo tanto el número de parámetros nulos es

$$I \cdot J + I \cdot K + I \cdot J \cdot K$$

pero, según mencionamos en el capítulo anterior,

$$\sum_{i=1}^{I} U_{12}(i,j) = \sum_{j=1}^{J} U_{12}(i,j) = 0$$

$$\sum_{i=1}^{I} U_{13}(i,k) = \sum_{k=1}^{K} U_{13}(i,k) = 0$$

$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

y por lo tanto, en este modelo, los parámetros nulos independientes son

$$(I-1)\cdot(J-1) + (I-1)\cdot(K-1) + (I-1)\cdot(J-1)\cdot(K-1).$$

Para el modelo $(12, 13)$ tenemos que

$$U_{23}(j,k) = 0 \ para \ j = 1,\ldots,J \ y \ para \ k = 1,\ldots,K$$

$$U_{123}(i,j,k) = 0 \ para \ i = 1,\ldots,I \ , \ para \ j = 1,\ldots,J \ y \ para \ k = 1,\ldots,K$$

y por lo tanto el número de parámetros nulos es

$$J \cdot K + I \cdot J \cdot K$$

pero, según mencionamos en el capítulo anterior,

$$\sum_{j=1}^{J} U_{23}(j,k) = \sum_{k=1}^{K} U_{23}(j,k) = 0$$

$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

y por lo tanto, en este modelo, los parámetros nulos independientes son

$$(J-1)\cdot(K-1) + (I-1)\cdot(J-1)\cdot(K-1).$$

Para el modelo $(12,23)$ tenemos que

$$U_{13}(i,k) = 0 \ para \ i = 1,\ldots,I \ y \ para \ k = 1,\ldots,K$$

$$U_{123}(i,j,k) = 0 \ para \ i = 1,\ldots,I \ , \ para \ j = 1,\ldots,J \ y \ para \ k = 1,\ldots,K$$

y por lo tanto el número de parámetros nulos es

$$I \cdot K + I \cdot J \cdot K$$

pero, según mencionamos en el capítulo anterior,

$$\sum_{i=1}^{I} U_{13}(i,k) = \sum_{k=1}^{K} U_{13}(i,k) = 0$$

$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

y por lo tanto, en este modelo, los parámetros nulos independientes son

$$(I-1)\cdot(K-1) + (I-1)\cdot(J-1)\cdot(K-1).$$

Para el modelo $(13,23)$ tenemos que

$$\sum_{i=1}^{I} U_{12}(i,j) = \sum_{j=1}^{J} U_{12}(i,j) = 0$$

$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

y por lo tanto el número de parámetros nulos es

$$I \cdot J + I \cdot J \cdot K$$

pero, según mencionamos en el capítulo anterior,

$$\sum_{i=1}^{I} U_{12}(i,j) = \sum_{j=1}^{J} U_{12}(i,j) = 0$$

$$\sum_{i=1}^{I} U_{123}(i,j,k) = \sum_{j=1}^{J} U_{123}(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}(i,j,k) = 0$$

y por lo tanto, en este modelo, los parámetros nulos independientes son

$$(I-1)\cdot(J-1) + (I-1)\cdot(J-1)\cdot(K-1).$$

Para el modelo (12, 13, 23) tenemos que

$$U_{123}\,(i,j,k) = 0 \ \ para \ \ i = 1,\ldots,I \ \ , \ \ para \ \ j = 1,\ldots,J \ \ y \ \ para \ \ k = 1,\ldots,K$$

y por lo tanto el número de parámetros nulos es

$$I \cdot J \cdot K$$

pero, según mencionamos en el capítulo anterior,

$$\sum_{i=1}^{I} U_{123}\,(i,j,k) = \sum_{j=1}^{J} U_{123}\,(i,j,k) = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} U_{123}\,(i,j,k) = 0$$

y por lo tanto, en este modelo, los parámetros nulos independientes son

$$(I-1)\cdot(J-1)\cdot(K-1)\,.$$

Para el modelo saturado tenemos que el número de parámetros nulos es cero.

APÉNDICE C

Correspondiente al modelo (1,2,3) tenemos

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln(O_{ijk})$$

$= O_{111}\ln(O_{111})+O_{112}\ln(O_{112})+O_{121}\ln(O_{121})+O_{122}\ln(O_{122})+O_{211}\ln(O_{211})+O_{212}\ln(O_{212})+ O_{221}\ln(O_{221})+O_{222}\ln(O_{222})$

$= 19\ln(19)+132\ln(132)+0\ln(0)+9\ln(9)+11\ln(11)+52\ln(52)+6\ln(6)+97\ln(97)$

$= 19\cdot 2.9444+132\cdot 4.8828+0\cdot(-\infty)+9\cdot 2.1972+11\cdot 2.3979+52\cdot 3.9512+6\cdot 1.7918+97\cdot 4.5747$

$= 55.9443+644.5299+0+19.7750+26.3768+205.4647+10.7506+443.7470=1406.5882$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111}\ln\left(\tilde{E}_{111}\right)+O_{112}\ln\left(\tilde{E}_{112}\right)+O_{121}\ln\left(\tilde{E}_{121}\right)+O_{122}\ln\left(\tilde{E}_{122}\right)+O_{211}\ln\left(\tilde{E}_{211}\right)+O_{212}\ln\left(\tilde{E}_{212}\right)+ O_{221}\ln\left(\tilde{E}_{221}\right)+O_{222}\ln\left(\tilde{E}_{222}\right)$

$= 19\ln(11.5985)+132\ln(93.4322)+0\ln(6.0702)+9\ln(48.8991)+11\ln(12.0334)+52\ln(96.9359)+ 6\ln(6.2979)+97\ln(50.7328)$

$= 19\cdot 2.4509+132\cdot 4.5372+0\cdot 1.8034+9\cdot 3.8898+11\cdot 2.4877+52\cdot 4.5740+6\cdot 1.8402+97\cdot 3.9266$

$= 46.5666+598.9152+0+35.0078+27.3645+237.8506+11.0413+380.8775=1337.6235$

$$G^2=-2\cdot(1337.6235-1406.5882)=137.9294$$

$gl=(I-1)\cdot(J-1)+(I-1)\cdot(K-1)+(J-1)\cdot(K-1)+(I-1)\cdot(J-1)\cdot(K-1)$

$=(2-1)\cdot(2-1)+(2-1)\cdot(2-1)+(2-1)\cdot(2-1)+(2-1)\cdot(2-1)\cdot(2-1)$

$=1\cdot 1+1\cdot 1+1\cdot 1+1\cdot 1\cdot 1=1+1+1+1=4$

Correspondiente al modelo (12,3) tenemos

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln(O_{ijk})=1406.5882$$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111}\ln\left(\tilde{E}_{111}\right)+O_{112}\ln\left(\tilde{E}_{112}\right)+O_{121}\ln\left(\tilde{E}_{121}\right)+O_{122}\ln\left(\tilde{E}_{122}\right)+O_{211}\ln\left(\tilde{E}_{211}\right)+O_{212}\ln\left(\tilde{E}_{212}\right)+ O_{221}\ln\left(\tilde{E}_{221}\right)+O_{222}\ln\left(\tilde{E}_{222}\right)$

$= 19 \ln(16.6748) + 132 \ln(134.3252) + 0 \ln(0.9939) + 9 \ln(8.0061) + 11 \ln(6.9571) + 52 \ln(56.0409) + 6 \ln(11.3742) + 97 \ln(91.6258)$

$= 19 \cdot 2.8139 + 132 \cdot 4.9003 + 0 \cdot (-0.0061) + 9 \cdot 2.0802 + 11 \cdot 1.9398 + 52 \cdot 4.0261 + 6 \cdot 2.4313 + 97 \cdot 4.5177$

$= 53.4641 + 646.8348 + 0 + 18.7218 + 21.3374 + 209.3563 + 14.5881 + 438.2182 = 1402.5206$

$$G^2 = -2 \cdot (1402.5206 - 1406.5882) = 8.1352$$

$gl = (I-1) \cdot (K-1) + (J-1) \cdot (K-1) + (I-1) \cdot (J-1) \cdot (K-1)$

$= (2-1) \cdot (2-1) + (2-1) \cdot (2-1) + (2-1) \cdot (2-1) \cdot (2-1) = 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 \cdot 1$

$= 1 + 1 + 1 = 3.$

Correspondiente al modelo (13, 2) tenemos

$$\sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} O_{ijk} \ln(O_{ijk}) = 1406.5882$$

$$\sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} O_{ijk} \ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111} \ln\left(\tilde{E}_{111}\right) + O_{112} \ln\left(\tilde{E}_{112}\right) + O_{121} \ln\left(\tilde{E}_{121}\right) + O_{122} \ln\left(\tilde{E}_{122}\right) + O_{211} \ln\left(\tilde{E}_{211}\right) + O_{212} \ln\left(\tilde{E}_{212}\right) + O_{221} \ln\left(\tilde{E}_{221}\right) + O_{222} \ln\left(\tilde{E}_{222}\right)$

$= 19 \ln(12.4724) + 132 \ln(92.5583) + 0 \ln(6.5276) + 9 \ln(48.4417) + 11 \ln(11.1595) + 52 \ln(97.8098) + 6 \ln(5.8405) + 97 \ln(51.1902)$

$= 19 \cdot 2.5235 + 132 \cdot 4.5278 + 0 \cdot 1.8760 + 9 \cdot 3.8804 + 11 \cdot 2.4123 + 52 \cdot 4.5830 + 6 \cdot 1.7648 + 97 \cdot 3.9355$

$= 47.9468 + 597.6747 + 0 + 34.9232 + 26.5352 + 238.3173 + 10.5889 + 381.7482 = 1337.7344$

$$G^2 = -2 \cdot (1337.7344 - 1406.5882) = 137.7096$$

$gl = (I-1) \cdot (J-1) + (J-1) \cdot (K-1) + (I-1) \cdot (J-1) \cdot (K-1)$

$= (2-1) \cdot (2-1) + (2-1) \cdot (2-1) + (2-1) \cdot (2-1) \cdot (2-1) = 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 \cdot 1$
$= 1 + 1 + 1 = 3.$

Correspondiente al modelo $(23,1)$ tenemos

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(O_{ijk}\right) = 1406.5882$$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111}\ln\left(\tilde{E}_{111}\right)+O_{112}\ln\left(\tilde{E}_{112}\right)+O_{121}\ln\left(\tilde{E}_{121}\right)+O_{122}\ln\left(\tilde{E}_{122}\right)+O_{211}\ln\left(\tilde{E}_{211}\right)+O_{212}\ln\left(\tilde{E}_{212}\right)+$
$O_{221}\ln\left(\tilde{E}_{221}\right) + O_{222}\ln\left(\tilde{E}_{222}\right)$

$= 19\ln(14.7239)+132\ln(90.3067)+0\ln(2.9448)+9\ln(52.0245)+11\ln(15.2761)+52\ln(93.6933)+$
$6\ln(3.0552) + 97\ln(53.9755)$

$= 19\cdot 2.6895+132\cdot 4.5032+0\cdot 1.0800+9\cdot 3.9517+11\cdot 2.7263+52\cdot 4.5400+6\cdot 1.1168+97\cdot 3.9885$

$= 51.1000 + 594.4239 + 0 + 35.5654 + 29.9892 + 236.0814 + 6.7011 + 386.8874 = 1340.7484$

$$G^2 = -2\cdot(1340.7484 - 1406.5882) = 131.6796$$

$gl = (I-1)\cdot(J-1) + (I-1)\cdot(K-1) + (I-1)\cdot(J-1)\cdot(K-1)$

$= (2-1)\cdot(2-1) + (2-1)\cdot(2-1) + (2-1)\cdot(2-1)\cdot(2-1) = 1\cdot 1 + 1\cdot 1 + 1\cdot 1\cdot 1$
$= 1+1+1 = 3.$

Correspondiente al modelo $(12,13)$ tenemos

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(O_{ijk}\right) = 1406.5882$$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111}\ln\left(\tilde{E}_{111}\right)+O_{112}\ln\left(\tilde{E}_{112}\right)+O_{121}\ln\left(\tilde{E}_{121}\right)+O_{122}\ln\left(\tilde{E}_{122}\right)+O_{211}\ln\left(\tilde{E}_{211}\right)+O_{212}\ln\left(\tilde{E}_{212}\right)+$
$O_{221}\ln\left(\tilde{E}_{221}\right) + O_{222}\ln\left(\tilde{E}_{222}\right)$

$= 19\ln(17.9312)+132\ln(133.0687)+0\ln(1.0688)+9\ln(7.9313)+11\ln(6.4518)+52\ln(56.5482)+$
$6\ln(10.5482) + 97\ln(92.4518)$

$= 19\cdot 2.8865+132\cdot 4.8909+0\cdot 0.0665+9\cdot 2.0708+11\cdot 1.8644+52\cdot 4.0351+6\cdot 2.3560+97\cdot 4.5267$

$= 54.8443 + 645.5943 + 0 + 18.6374 + 20.5080 + 209.8249 + 14.1357 + 439.0887 = 1402.6331$

$$G^2 = -2\cdot(1402.6331 - 1406.5882) = 7.9102$$

$gl = (J-1)\cdot(K-1) + (I-1)\cdot(J-1)\cdot(K-1) = (2-1)\cdot(2-1) + (2-1)\cdot(2-1)\cdot(2-1)$
$= 1\cdot 1 + 1\cdot 1\cdot 1 = 1+1 = 2.$

Correspondiente al modelo $(12,23)$ tenemos

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(O_{ijk}\right) = 1406.5882$$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111}\ln\left(\tilde{E}_{111}\right)+O_{112}\ln\left(\tilde{E}_{112}\right)+O_{121}\ln\left(\tilde{E}_{121}\right)+O_{122}\ln\left(\tilde{E}_{122}\right)+O_{211}\ln\left(\tilde{E}_{211}\right)+O_{212}\ln\left(\tilde{E}_{212}\right)+$ $O_{221}\ln\left(\tilde{E}_{221}\right)+O_{222}\ln\left(\tilde{E}_{222}\right)$

$= 19\ln(21.1682)+132\ln(129.8318)+0\ln(0.4821)+9\ln(8.5179)+11\ln(8.8318)+52\ln(54.1682)+$ $6\ln(5.5179)+97\ln(97.4821)$

$= 19\cdot 3.0525+132\cdot 4.8662+0\cdot(-0.7296)+9\cdot 2.1422+11\cdot 2.1784+52\cdot 3.9921+6\cdot 1.7080+97\cdot 4.5797$

$= 57.9975+642.3436+0+19.2795+23.9619+207.5889+10.2480+444.2279 = 1405.6474$

$$G^2 = -2\cdot(1405.6474-1406.5882) = 1.8816$$

$gl = (I-1)\cdot(K-1)+(I-1)\cdot(J-1)\cdot(K-1) = (2-1)\cdot(2-1)+(2-1)\cdot(2-1)\cdot(2-1)$ $= 1\cdot 1+1\cdot 1\cdot 1 = 1+1 = 2.$

Correspondiente al modelo $(13,23)$ tenemos

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(O_{ijk}\right) = 1406.5882$$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111}\ln\left(\tilde{E}_{111}\right)+O_{112}\ln\left(\tilde{E}_{112}\right)+O_{121}\ln\left(\tilde{E}_{121}\right)+O_{122}\ln\left(\tilde{E}_{122}\right)+O_{211}\ln\left(\tilde{E}_{211}\right)+O_{212}\ln\left(\tilde{E}_{212}\right)+$ $O_{221}\ln\left(\tilde{E}_{221}\right)+O_{222}\ln\left(\tilde{E}_{222}\right)$

$= 19\ln(15.8333)+132\ln(89.4621)+0\ln(3.1667)+9\ln(51.5379)+11\ln(14.1667)+52\ln(94.5379)+$ $6\ln(2.8333)+97\ln(54.4621)$

$= 19\cdot 2.7621+132\cdot 4.4938+0\cdot 1.1527+9\cdot 3.9423+11\cdot 2.6509+52\cdot 4.5490+6\cdot 1.0414+97\cdot 3.9975$

$= 52.4802+593.1836+0+35.4809+29.1598+236.5480+6.2487+387.7580 = 1340.8592$

$$G^2 = -2\cdot(1340.8592-1406.5882) = 131.4580$$

$gl = (I-1)\cdot(J-1)+(I-1)\cdot(J-1)\cdot(K-1) = (2-1)\cdot(2-1)+(2-1)\cdot(2-1)\cdot(2-1)$ $= 1\cdot 1+1\cdot 1\cdot 1 = 1+1 = 2.$

Correspondiente al modelo $(12,13,23)$ tenemos

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(O_{ijk}\right)=1406.5882$$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111}\ln\left(\tilde{E}_{111}\right)+O_{112}\ln\left(\tilde{E}_{112}\right)+O_{121}\ln\left(\tilde{E}_{121}\right)+O_{122}\ln\left(\tilde{E}_{122}\right)+O_{211}\ln\left(\tilde{E}_{211}\right)+O_{212}\ln\left(\tilde{E}_{212}\right)+ O_{221}\ln\left(\tilde{E}_{221}\right)+O_{222}\ln\left(\tilde{E}_{222}\right)$

$= 19\ln(18.6745)+132\ln(132.3254)+0\ln(0.3257)+9\ln(8.6744)+11\ln(11.3255)+52\ln(51.6746)+ 6\ln(5.6743)+97\ln(97.3256)$

$= 19\cdot 2.9272+132\cdot 4.8853+0\cdot(-1.1218)+9\cdot 2.1604+11\cdot 2.4271+52\cdot 3.9450+6\cdot 1.7359+97\cdot 4.5781$

$= 55.6160+644.8549+0+19.4434+26.6976+205.1383+10.4157+444.0720=1406.2378$

$$G^2=-2\cdot(1406.2378-1406.5882)=0.7008$$

$gl=(I-1)\cdot(J-1)\cdot(K-1)=(2-1)\cdot(2-1)\cdot(2-1)=1\cdot 1\cdot 1=1.$

Correspondiente al modelo saturado tenemos

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(O_{ijk}\right)=1406.5882$$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2} O_{ijk}\ln\left(\tilde{E}_{ijk}\right)$$

$= O_{111}\ln\left(\tilde{E}_{111}\right)+O_{112}\ln\left(\tilde{E}_{112}\right)+O_{121}\ln\left(\tilde{E}_{121}\right)+O_{122}\ln\left(\tilde{E}_{122}\right)+O_{211}\ln\left(\tilde{E}_{211}\right)+O_{212}\ln\left(\tilde{E}_{212}\right)+ O_{221}\ln\left(\tilde{E}_{221}\right)+O_{222}\ln\left(\tilde{E}_{222}\right)$

$= 19\ln(19)+132\ln(132)+0\ln(0)+9\ln(9)+11\ln(11)+52\ln(52)+6\ln(6)+97\ln(97)$

$= 19\cdot 2.9444+132\cdot 4.8828+0\cdot(-\infty)+9\cdot 2.1972+11\cdot 2.3979+52\cdot 3.9512+6\cdot 1.7918+97\cdot 4.5747$

$= 55.9443+644.5299+0+19.7750+26.3768+205.4647+10.7506+443.7470=1406.5882$

$$G^2=-2\cdot(1406.5882-1406.5882)=0$$

$$gl=0.$$