

La Prueba Q de Cochran. Teoría y aplicaciones en Biología

Resumen

Se muestra el planteamiento estadístico y las aplicaciones de la prueba Q de Cochran para muestras grandes. Esta técnica estadística perteneciente a la "familia" de análisis de ji cuadrada para datos enumerativos, se basa en un diseño en bloques completos aleatorizados. El enfoque de presentación sigue el esquema del artículo original de Cochran (1950). Primeramente se describe la prueba de McNemar para muestras pareadas, y posteriormente la prueba Q como una generalización de aquella. Al deducir el estadístico Q es posible revisar varios conceptos estadísticos básicos, como los de probabilidad condicional, valor esperado, varianza y covarianza de variables aleatorias. También se ilustra la aplicación de la prueba Q y de las pruebas de comparación múltiple *post hoc*, con datos de un estudio realizado por Rico-Gray (1993), para analizar el uso de fuentes alimenticias por hormigas. Por último se remarcan las suposiciones del modelo estadístico para Q y se mencionan dos generalizaciones de esta prueba.

1 Introducción

El diseño y análisis estadístico de experimentos proporciona una colección de métodos para la obtención de datos y su aplicación en la estimación de parámetros y en las pruebas de significación, principalmente para la comparación de tratamientos considerados en una investigación. La aplicación particular de cada método depende, claro, de la situación experimental seleccionada, bajo el contexto del objetivo de la investigación. Es en el trabajo interdisciplinario en donde se pueden conjuntar las diferentes perspectivas de investigación, y de éstas, el estadístico puede tener una contribución relevante.

Estas notas son el resultado de la interacción académica con biólogos de diferentes áreas: ecólogos, biólogos marinos y botánicos. Los conjuntos de datos que deseaban analizar ellos, tenían algo en común: consideraban la comparación de muestras o grupos

Jorge A. Navarro Alberto
Universidad Autónoma de Yucatán
México

experimentales con base en variables dicotómicas. Las investigaciones demandaban la respuesta a la siguiente pregunta: ¿Qué método estadístico sería adecuado para comparar grupos ("tratamientos" o "muestras") sobre los cuales se registran valores de una variable de respuesta (discreta) dicotómica (es decir, aquella que toma solamente uno de dos posibles valores) y con la propiedad de que los grupos están apareados, formando bloques? A primera vista, el enfoque no debía ser el mismo que el aplicado para cuando la variable de respuesta fuera continua. W. G. Cochran (1950) dio respuesta a la pregunta, al proponer un método, la prueba Q , para hacer la comparación de dos o más tratamientos sobre datos enumerativos cuando la variable de respuesta es dicotómica y el muestreo se realiza de acuerdo a un diseño en bloques completos al azar.

El objetivo de estas notas es mostrar el planteamiento estadístico y las aplicaciones de la prueba Q de Cochran para muestras grandes, resaltando su sencillez, ventajas y limitaciones en el trabajo práctico del biólogo. Estas técnicas estadísticas se ubican dentro de la "familia" de análisis ji cuadrada para datos enumerativos. He escogido el enfoque de presentación (no muy riguroso) del artículo original de Cochran (1950), pues es intuitivo y accesible para los estudiantes de ciencias biológicas. Blomqvist (1951) dio una prueba de los resultados asintóticos respecto a la distribución de Q , pero requiere conocimientos profundos de la estadística matemática.

2 El diseño en bloques completos aleatorizados

Uno de los diseños más empleados en las ciencias naturales e ingeniería, es el de bloques completos aleatorizados. Al procurar en este diseño que las unidades experimentales homogéneas (los bloques) reciban todos los tratamientos, se busca mejorar la precisión de los experimentos en dos aspectos. Primero, eliminar la variación producida por la heterogeneidad que pudieran tener las unidades de manera "natural" pero incontrolable (aunque la eliminación de la variación no sea total). Segundo, eliminar la variabilidad introducida por los investigadores en el desarrollo mismo del experimento. En consecuencia, hay una disminución sustancial del error experimental, con la consiguiente ventaja de poder detectar mejor los efectos de los tratamientos (Montgomery, 1991).

Para el análisis de resultados experimentales, es común recurrir a pruebas de significación estadística. Así, en la literatura se remarca que la correcta aplicación de alguna prueba de la "familia de pruebas t ", depende de supuestos de normalidad de ciertas variables (ver párrafo siguiente). Análogamente, cuando se requiere el uso de la prueba F , debe satisfacerse el supuesto de homogeneidad de varianzas. Dado que estos supuestos usualmente se satisfacen cuando la variable de respuesta es continua, son más conocidos los casos del análisis de diseños en bloques, cuando las estimaciones y pruebas se hacen sobre variables de respuesta continuas medidas en las unidades experimentales.

Por ejemplo, se aplica la prueba t de muestras apareadas cuando hay dos tratamientos o grupos experimentales; en esta prueba, la suposición es que las diferencias $d_j = y_{1j} - y_{2j}$ de cada pareja de valores de la variable de respuesta, tengan una distribución aproximadamente normal. O bien, cuando hay más de dos tratamientos, se aplica la prueba F de Análisis de Varianza, ANDEVA) para un diseño en bloques. (Zar, 1984, págs. 150-153; 222-226). También es conocido que tanto las pruebas t como F , son lo suficientemente robustas contra la no normalidad cuando las muestras son grandes. No obstante, cuando la variable de respuesta es dicotómica, la forma en que podrían

compararse dos tratamientos “por parejas” puede enfocarse de manera distinta, sin la preocupación del cumplimiento de los supuestos de normalidad o de homogeneidad de varianzas. El ejemplo de la siguiente sección ilustra la aplicación de la conocida prueba de McNemar para datos enumerativos, basados en diseños por parejas (o “pareados”).

3 Diseños comparativos por parejas

Ejemplo 1

Supóngase que se quieren comparar dos líneas X y Y de un virus que ataca a las hojas de una especie de planta. (Fig. 1). El experimento tratará de estimar la capacidad de infección que tienen las líneas y, para ello, se inoculan una a cada lado de la vena central (o costa) de las hojas seleccionadas. (Esto funcionará siempre y cuando los dos focos de infección no comiencen a interactuar.)

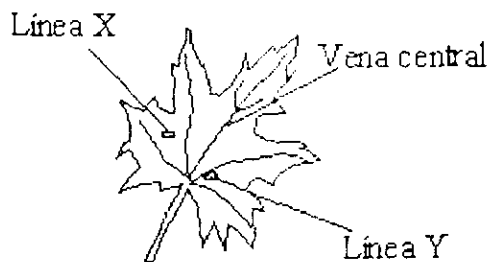


Figura 1 Ejemplo de un diseño comparativo por parejas. Se inoculan las líneas X y Y de un virus, en cada lado de la vena central de la hoja.

Con este procedimiento, las diferentes sensibilidades a la inoculación que pueden caracterizar a hojas diferentes, quedan eliminadas y se evalúan únicamente los méritos relativos de los dos tratamientos. Si éstos se comparan respecto a una variable continua, entonces una prueba disponible es la “t de muestras pareadas” (siempre que se satisfagan los supuestos de normalidad de las diferencias calculadas en la prueba.) Por ejemplo, podrían compararse las dos líneas respecto a la media de los tiempos entre la inoculación y la manifestación de la enfermedad provocada por el virus, calculándose el estadístico *t* y comparándose con valores críticos asociados a un nivel de significación preestablecido (consúltese Sincich (1993), págs. 487-490).

Pero hay casos en donde lo que recolectamos son datos enumerativos. Por ejemplo, podríamos preguntarnos si las hojas se infectan más frecuentemente con los virus de la línea X que con los de la línea Y y, para ello, observamos la reacción positiva o negativa de las hojas a las líneas. Aquí también podemos usar como parte del diseño, el método de aparear los datos, inoculando en las “mitades” de las hojas, de modo que cada hoja se convierta en un bloque. Las líneas X y Y generarán muestras pareadas.

		Línea Y		
		0	1	Total
Línea X	0	25 (a)	12 (b)	37
	1	33 (c)	45 (d)	38
Total		58	57	<i>n</i> = 115

Tabla 1 Datos ficticios para ilustrar la comparación de proporciones poblacionales, a partir de dos muestras.

La tabla de frecuencias 2×2 (Tabla 1) clasifica las reacciones positivas (codificadas como 1's) o negativas (codificadas como 0's) a la inoculación en hojas de las líneas X y Y , a partir del diseño de muestras pareadas. Los datos son ficticios y se suponen obtenidos de $n = 115$ hojas.

El pareamiento ocurre porque cada combinación (0,0), (0,1), (1,0) y (1,1) proviene de la observación a la reacción respecto a las líneas X y Y , respectivamente, en la *misma* hoja. Por tanto, las muestras generadas por la línea X contienen exactamente las mismas hojas que las muestras generadas por la línea Y .

Análisis estadístico

Como señala Cochran (1950), "debido a que el apareamiento puede introducir correlación entre los resultados de las muestras distintas, invalida la prueba ji cuadrada ordinaria, lo cual da muy pocos resultados significativos cuando el pareamiento es efectivo." McNemar (1947) encontró el procedimiento apropiado para probar la hipótesis de que las respuestas nulas (0's) se presentan en la población en igual proporción para ambos "tratamientos", tomando en cuenta el pareamiento de las muestras. En nuestro ejemplo, la hipótesis nula a probar es que la proporción poblacional de hojas no infectadas (0's) con la línea X es la misma proporción poblacional de las hojas infectadas con la línea Y .

Ejemplo 1 (continuación)

La idea de la prueba de McNemar es comparar qué tan iguales son las frecuencias relativas $\frac{a+b}{n} = \frac{37}{115}$ y $\frac{a+c}{n} = \frac{58}{115}$, donde $a = 25$, $b = 12$ y $c = 33$ son los números observados de (0,0), (0,1) y (1,0) respectivamente y $n = 115$ es el total de muestras ($d = 45$ es el número observado de (1,1)'s. Todos estos datos aparecen en la Tabla 1.) La hipótesis nula establece que la proporción de reacciones negativas que se dieron en las hojas, es la misma para ambas líneas. La prueba se deduce entonces notando que ambos numeradores $a+c$ y $a+b$ contienen a las $a = 25$ hojas que tuvieron la misma reacción negativa (0,0). De aquí que la igualdad de los numeradores implicaría que el número de hojas que reaccionaron negativamente a la línea X y positivamente a la línea Y (los resultados de la forma (0,1)) es el mismo número de hojas que reaccionaron positivamente a la línea X y negativamente a la línea Y (los resultados de la forma (1,0)). En otras palabras, si ambas líneas infectan de igual modo a las hojas, esperaríamos la mitad del número de hojas con reacciones (0,1) y la mitad con reacciones (1,0). Luego, la prueba puede hacerse probando si los números b y c son éxitos y fracasos binomiales de $m = b + c$ ensayos con probabilidad $\frac{1}{2}$. Las frecuencias, esperadas para (0,1) y (1,0) serán entonces iguales a $\frac{1}{2} m$. Por ello, el estadístico que prueba tales frecuencias observadas con las esperadas es:

$$\chi^2 = \frac{\left(b - \frac{1}{2} m\right)^2}{\frac{1}{2} m} + \frac{\left(c - \frac{1}{2} m\right)^2}{\frac{1}{2} m} = \frac{(b - c)^2}{b + c} \quad \text{con 1 grado de libertad.}$$

Una corrección para continuidad puede aplicarse, del siguiente modo:

$$\chi_c^2 = \frac{(|b - c| - 1)^2}{b - c}$$

En este ejemplo, $b = 33$, $c = 12$, por lo que para $\alpha = .05$ y:

$$\chi_c^2 = \frac{(|33 - 12| - 1)^2}{33 + 12} = 8.89$$

rechazamos la hipótesis nula de que las proporciones de no-infección es la misma para las líneas X y Y , $\chi_{0.05,1}^2 = 3.841$, $0.001 < P < 0.005$.

4 La prueba Q de Cochran como generalización de la prueba de McNemar

La prueba de McNemar se extiende a situaciones en las que hay más de dos muestras. Un ejemplo lo da el siguiente estudio sobre el uso de recursos alimenticios de organismos.

Ejemplo 2

Rico-Gray (1993) estudió el uso de recursos alimenticios por especies de hormigas, las cuales "seleccionan" fuentes alimenticias líquido/energía en la vegetación de la planicie tropical de la costa de Veracruz, México. Las fuentes alimenticias clasificadas por Rico-Gray fueron ocho: FL (néctar floral), NER (néctar de otras estructuras reproductivas), E (elaiosoma), HOM (*Homoptera*), LEP (*Lepidoptera*), NEF (néctar extrafloral), FL/S (flores sobre el suelo) y FR/S (frutos sobre el suelo), y de estas ocho, FL, NER, HOM y NEF fueron las fuentes más utilizadas. Un problema de interés ecológico es responder si hay diferencia o no respecto a la preferencia de alguna de estas últimas cuatro fuentes, por parte de las especies de hormiga.

Estadísticamente se puede trasladar este problema a probar que las fuentes alimenticias tienen igual porcentaje de selección (por parte de las especies de hormiga). Aquí hay más de dos muestras, pues hay 4 fuentes utilizables y, de acuerdo con la descripción de cómo se hizo la investigación, saldrá a relucir la aplicación del principio de bloqueo.

Método

En un tramo de 1 km del área de estudio, Rico-Gray hizo observaciones de campo durante un año en visitas quincenales (que requerían tres días en total por visita), en las cuales registró todas las ocurrencias de hormigas que estuvieran colectando líquidos de diferentes especies de plantas o que estuvieran en contacto con otros insectos que se alimentan del néctar de plantas (como algunos homópteros o lepidópteros), o que estuvieran libando de flores o de frutos caídos. Las ocurrencias de las especies de hormigas se hicieron en plantas que visitaron todas ellas durante todo el estudio, y que fueron marcadas para su revisión en cada observación de campo. Las unidades de muestreo fueron las gotas de néctar cuya fuente era alguna estructura que proporciona néctar

como, por ejemplo, una planta de las plantas marcadas, o de algún nectarífigo asociado a éstas. El análisis Q compararía las proporciones de uso de las 4 fuentes más usadas.

Se consideró que una especie de hormiga tenía preferencia por una determinada fuente alimenticia (es decir, se registraba: "usa la fuente"), cuando individuos de la especie en cuestión estaban inmóviles, con sus partes bucales en contacto con los tejidos secretores de néctar o estuvieran palpando con su antena los depósitos de néctar de los Homópteros o Lepidópteros portadores, durante varios minutos. Si esto no sucedía en ninguna de las observaciones de campo, se registraba "no usa de la fuente."

Especies de hormiga	Fuente alimenticia			
	FL	NER	HOM	NEF
1) <i>Azteca</i> sp.	1	1	1	0
2) <i>Brachymyrex</i> sp.	0	1	1	0
3) <i>Camponotus mucronatus</i> Emery	1	1	1	1
4) <i>Camponotus planatus</i> Roger	1	1	1	1
5) <i>Camponotus rectangularis</i> Emery	0	1	1	0
6) <i>Camponotus sereiventris</i> (Guerin)	0	1	1	0
7) <i>Camponotus</i> (<i>Myrmobrachys</i>) sp.	1	1	1	1
8) <i>Camponotus</i> sp.	1	1	1	1
9) <i>Cronomyrma</i> sp.	1	1	1	1
10) <i>Crematogaster brevispinosa</i> Mayr	1	1	1	1
11) <i>Forelius</i> sp.	1	1	1	1
12) <i>Monomorium</i> sp.	1	1	1	1
13) <i>Paratrechina longicornis</i> (Cantreille)	1	1	1	0
14) <i>Pheidole</i> sp.	1	1	1	0
15) <i>Pseudomyrmex ejectus</i> (Smith)	1	1	0	1
16) <i>Pseudomyrmex ferrugineus</i> (Smith)	0	0	0	1
17) <i>Pseudomyrmex filiformis</i> (Fabricius)	1	1	1	1
18) <i>Pseudomyrmex gracilis</i> (Fabricius)	1	1	1	1
19) <i>Pseudomyrmex ita</i> (Forel)	1	0	0	0
20) <i>Pseudomyrmex pallidus</i> (Smith)	0	1	1	1
21) <i>Pseudomyrmex simplex</i> (Smith)	0	1	1	0
22) <i>Solenopsis geminata</i> (Fabricius)	1	0	1	0
23) <i>Zacryptocerus</i> sp.	1	1	1	0
24) Desconocida 1	1	0	0	0
25) Desconocida 2	0	1	0	0
26) Desconocida 3	1	0	0	0
27) Desconocida 4	0	0	0	0
28) Desconocida 5	0	0	0	0
29) Desconocida 6	0	0	0	0
30) Desconocida 7	0	0	0	0

Tabla 2 Uso de las cuatro fuentes alimenticias líquido/energía más frecuentadas por especies de hormigas en la vegetación de la zona baja tropical seca de la costa de Veracruz (estación biológica El Morro de la Mancha.) Los datos fueron obtenidos por Rico-Gray (1993). Los tipos de recursos alimenticios presentados aquí son: FL (néctar floral); NER (néctar de otras estructuras reproductivas); HOM (Homoptera) y NEF (néctar extrafloral.)

Debido a que en el área de estudio cada especie de hormiga difiere en abundancia y comportamiento (etología), biológicamente cada una tiene estrategias diferentes de acceso a las fuentes de néctar (específicamente, a las gotas de néctar que provienen de cada fuente), y por tanto, nos sugiere que las especies sean consideradas como bloques. Claro, aquí se supone que las cuatro fuentes alimenticias estaban disponibles para los individuos de todas las especies de hormiga registradas, para que al tener a las especies de hormiga como bloques, todos los "tratamientos" (fuentes alimenticias) estén representados en cada bloque.

Los resultados obtenidos por Rico-Gray aparecen en la Tabla 2. Esta presentación es la que parece más comprensible para mostrar las posibles combinaciones de resultados entre todas las muestras.

Hipótesis nula y alternativa

Al escribir las distintas combinaciones de 0's y 1's observadas, una combinación por renglón (como en la Tabla 2), podemos representar la forma general de la tabla preparada para el análisis estadístico, como aparece en la Tabla 3.

Bloques	Muestras (grupos)				Totales
	1	2	...	c	
1	y_{11}	y_{12}	...	y_{1c}	B_1
2	y_{21}	y_{22}	...	y_{2c}	B_2
...
r	y_{r1}	y_{r2}	...	y_{rc}	B_r
Totales	G_1	G_2	...	G_c	$\Sigma G_j = \Sigma B_i$

Tabla 3 Forma general de presentación de datos dicotómicos ($y_{ij} = 0$ ó $y_{ij} = 1$) para la prueba Q de Cochran.

Aquí y_{ij} solamente puede tomar dos valores: 0 o 1. Supóngase que r , el número de renglones (bloques) es grande. La probabilidad de un 1 se supone que varía de renglón a renglón (en el ejemplo, recuérdese que las hormigas pueden tener diferentes estrategias para acceder a cada uno de los recursos). El número total observado de éxitos (1's) en el i -ésimo renglón (bloque), denotado por B_i , es la suma de los 1's de ese renglón; su valor se considera fijo.

En general, la hipótesis nula a probar y la hipótesis alternativa, se pueden establecer "coloquialmente" como:

H_0 : Los grupos son homogéneos.

H_A : Hay diferencia en los grupos.

Sean $p_{ij} = P(y_{ij} = 1)$, $i = 1, \dots, r$, $j = 1, \dots, c$ las probabilidades poblacionales de ocurrencia de un éxito para cada "celda" (i, j), entonces, la homogeneidad de los grupos implica que:

$$p_{i1} = p_{i2} = \dots = p_{ic} \text{ para cada } i \text{ de } 1 \text{ a } r.$$

Esto es, para cada bloque, la probabilidad de un éxito en un grupo no depende de qué grupo se trata específicamente. Entonces, las hipótesis pueden ser establecidas de nuevo (Conover 1971):

$$H_0: p_{i1} = p_{i2} = \dots = p_{ic}, \text{ para cada } i, i = 1, \dots, r.$$

$$H_A: p_{ij} \neq p_{ik}, \text{ para algún } j \text{ y algún } k, \text{ y para algún } i.$$

Hacemos notar aquí que la comparación de proporciones poblacionales depende de una restricción: la presencia de bloques. Éstos se incorporan como un efecto (aleatorio) aditivo en el modelo estadístico lineal que describe a la variable de respuesta dicotómica y_{ij} . Esto lo remarca Conover (1971) al afirmar que, en este análisis, "los bloques son seleccionados aleatoriamente de las poblaciones de los posibles bloques." (Secc. 5, inciso 3).

Así, H_0 establece que las p_{ij} a lo largo de cada renglón son iguales entre columnas, pero pueden variar de un bloque (renglón) a otro.

Si es cierta la hipótesis nula, las c columnas se considerarán igualmente probables para obtener uno de los éxitos de los B_i que hay en el renglón i . La población de posibles resultados en el i -ésimo renglón consiste de las $\binom{c}{B_i}$ formas en que los B_i éxitos pueden distribuirse entre las c columnas.

Como sucede con la prueba ji cuadrada ordinaria (en donde las columnas son independientes), el criterio de prueba que propuso Cochran fue la consideración del cálculo de

$$SC_{grupos} = \sum_{j=1}^c (G_j - \bar{G})^2$$

donde G_j es el total de éxitos en la j -ésima columna (grupo) y \bar{G} es la media de los totales por grupo. La razón de elegir este estadístico se debe a que el uso de los datos de un renglón para la detección de diferencias entre columnas, puede depender de la probabilidad de éxito en ese renglón.

La meta que perseguimos es determinar si SC_{grupos} tiene alguna distribución de muestreo límite conocida, cuando el número de renglones, r , es grande. El siguiente razonamiento nos permite llegar a tal distribución límite.

El estadístico Q de Cochran para muestras grandes

El desarrollo para la obtención del estadístico para la prueba de igualdad de proporciones de éxitos entre grupos, inicia con el cálculo de las probabilidades de éxito y de fracaso en cada "celda" ubicada en el renglón i y columna j :

$$P(y_{ij} = 1) = \frac{B_i}{c} \quad \text{y} \quad P(y_{ij} = 0) = 1 - \frac{B_i}{c} \quad (1)$$

A partir de la definición de valor esperado y la varianza de una variable aleatoria, o bien, notando que y_{ij} es una variable aleatoria Bernoulli, con probabilidades de éxito y fracaso dadas por las fórmulas en (1), son inmediatas las comprobaciones de las siguientes igualdades:

$$E(y_{ij}) = \frac{B_i}{c} \quad (2)$$

$$\sigma^2 = \text{Var}(y_{ij}) = \frac{B_i}{c} \left(1 - \frac{B_i}{c}\right) \quad (3)$$

La covarianza de y_{ij} y y_{ik} es:

$$\text{Cov}(y_{ij}, y_{ik}) = E(y_{ij}, y_{ik}) - E(y_{ij}) E(y_{ik})$$

en donde $E(y_{ij}, y_{ik}) = \sum_{\text{todo } y_j, y_k} y_{ij}, y_{ik} p(y_{ij}, y_{ik})$. Aquí $p(y_{ij}, y_{ik})$ es la probabilidad

conjunta de y_{ij} y y_{ik} . Sin embargo, tanto y_{ij} como y_{ik} toman los valores 0 y 1, por lo que el único par que contribuye a un término no-nulo en $E(y_{ij}, y_{ik})$ es la pareja (1,1), pues en cualquier otro caso, el producto $y_{ij}y_{ik}$ que aparece en la suma contiene al menos un factor igual a 0. Por tanto, la probabilidad conjunta discreta de y_{ij} y y_{ik} que basta calcular es:

$$p(1,1) = p(1) p(1 | 1)$$

donde $p(y_{ij} | y_{ik})$ es la probabilidad condicional de y_{ij} dado y_{ik} . Por la fórmula (1) se tiene que

$$p(1) = \frac{B_i}{c}$$

y

$$p(1 | 1) = \frac{B_i - 1}{c - 1}$$

porquedespuésdepresentarseun 1, quedan $B_i - 1$ unos en el renglón i , del número de columnas que quedan, $c - 1$.

Así:

$$p(y_{ij}, y_{ik}) = \frac{B_i}{c} \left(\frac{B_i - 1}{c - 1}\right) = \frac{B_i(B_i - 1)}{c(c - 1)} \Rightarrow E(y_{ij}, y_{ik}) = 1 \cdot 1 \cdot p(1,1) = \frac{B_i(B_i - 1)}{c(c - 1)}$$

$$\text{Cov}(y_{ij}, y_{ik}) = E(y_{ij}, y_{ik}) - E(y_{ij}) E(y_{ik})$$

$$= \frac{B_i(B_i - 1)}{c(c - 1)} - \frac{B_i}{c} \cdot \frac{B_i}{c}$$

$$= \frac{-\frac{B_i}{c} \left(1 - \frac{B_i}{c}\right)}{c - 1}$$

La aleatorización es independiente en diferentes renglones, entonces $\text{Cov}(y_{ij}, y_{ij'}) = 0$. Por tanto, las medias, las varianzas y las covarianzas de los totales de columna G_j serán las sumas de las correspondientes medias, varianzas y covarianzas de las y_{ij} (Mendenhall *et al*, 1986).

$$E(G_j) = E\left(\sum_{i=1}^r y_{ij}\right) = \sum_{i=1}^r E(y_{ij}) = \sum_{i=1}^r \frac{B_i}{c}$$

$$\sigma^2(G_j) = \text{Var} \left(\sum_{i=1}^r y_{ij} \right) = \sum_{i=1}^r \sigma^2(y_{ij}) + 2 \sum_{i < l} \text{Cov}(y_{ij}, y_{il}) = \sum_{i=1}^r \sigma^2(y_{ij})$$

$$\sigma^2(G_j) = \sum_{i=1}^r \frac{B_i}{c} \left(1 - \frac{B_i}{c} \right)$$

$$\text{Cov}(G_j, G_k) = \sum_{i=1}^r \sum_{l=1}^r \text{Cov}(y_{ij}, y_{lk}) = \sum_{i=1}^r \text{Cov}(y_{ij}, y_{ik}) =$$

$$\sum_{i=1}^r \frac{\left(-\frac{B_i}{c} \left(1 - \frac{B_i}{c} \right) \right)}{c-1} = \frac{-1}{c-1} \sigma^2(G_j)$$

Si el número de renglones es grande, la distribución conjunta de G_1, G_2, \dots, G_c tenderá a la distribución normal multivariada. De acuerdo con un resultado de Walsh (1947), el hecho de que el conjunto de c variables G_j sigan una distribución normal multivariada con varianza común σ^2 y covarianza $\rho\sigma^2$, implica que $SC_{grupos} = \sum_{j=1}^c (G_j - \bar{G})^2$ se distribuye $\chi^2\sigma^2(1 - \rho)$, con $(c - 1)$ grados de libertad. En el caso que incumbe a G_j , denotemos por σ^2 a la varianza de los totales por grupo $\sigma^2(G_j)$, entonces:

$$\sigma^2 = \sum_{i=1}^r \frac{B_i}{c} \left(1 - \frac{B_i}{c} \right) \quad \rho\sigma^2 = \frac{-1}{c-1} \sigma^2$$

$$\Rightarrow \sigma^2(1 - \rho) = \sum_{i=1}^r \frac{B_i}{c} \left(1 - \frac{B_i}{c} \right) \left(1 + \frac{1}{c-1} \right)$$

$$= \frac{c}{c-1} \sum_{i=1}^r \frac{B_i}{c} \left(1 - \frac{B_i}{c} \right) = \frac{1}{c-1} \sum_{i=1}^r B_i \left(1 - \frac{B_i}{c} \right)$$

Por tanto, cuando r es grande:

$$Q = \frac{SC_{grupos}}{\sigma^2(1 - \rho)} = \frac{\sum_{j=1}^c (G_j - \bar{G})^2}{\frac{1}{c-1} \sum_{i=1}^r \left[B_i \left(1 - \frac{B_i}{c} \right) \right]}$$

$$= (c-1) \frac{\left[\sum_{j=1}^c G_j^2 - \left(\sum_{j=1}^c G_j \right)^2 \right]}{\sum_{i=1}^r B_i - \frac{\sum_{i=1}^r B_i^2}{c}}$$

se distribuye χ^2_{c-1} ; la última fórmula es la misma expresión que da Zar (1984), salvo que usa a y b en lugar de c y r , respectivamente.

Observaciones

1. El valor de Q (ecuación 4) se compara con el valor tabulado de χ^2 cuadrado con $c - 1$ grados de libertad. Esta aproximación a la χ^2 cuadrado es buena, siempre que $r \geq 4$ y $rc \geq 24$, como lo señalan empíricamente Tate y Brown (1970). De otro modo, se debería usar la distribución exacta de Q (Patil, 1975). Está disponible un programa que escribí en FORTRAN (llamado QCOCH), que lee datos de un archivo en formato ASCII para hacer el cálculo de Q y el nivel P de la prueba χ^2 cuadrado. También el programa ASP, que viene incluido en Sincich (1993) (programa producido por la compañía DMC Software), contiene la prueba Q de Cochran para muestras grandes. He escrito un programa en FORTRAN que calcula la distribución exacta de Q si $c = 3$, basado en el artículo de Patil (1975). Este autor no utiliza un enfoque permutacional clásico, como en la prueba exacta de Fisher para tablas de contingencia 2×2 , sino algorítmico (según el cumplimiento de una igualdad), para enumerar ciertos subconjuntos que se forman con las ternas de valores posibles de una observación, y con ello consigue de una manera relativamente sencilla, el cálculo de la probabilidad de ocurrencia de todos los valores posibles de Q para un tamaño de muestra fijo. Incluso, el algoritmo es posible generalizarlo a $c \geq 3$ grupos (formando c -uplas de valores posibles).
2. Nótese que si todos los elementos del renglón i son ceros, entonces $B_i = 0$; si eliminamos del análisis todos los renglones con ceros, el valor de Q no varía. Esto origina que el número de renglones (bloques) se reduzca.
3. Análogamente, si todos los elementos del renglón i en la Tabla 3 son 1's, entonces $B_i = c$. En la expresión (4) que define a Q , el denominador contiene al factor $(1 - B_i/c)$, el cual se anula para tal valor de B_i . Por tanto, todos los renglones que contienen exclusivamente 1's no contribuyen en una cantidad no nula (positiva) al valor del denominador de Q . También es fácil ver que tampoco afectan el valor del numerador de Q . (Como bien se sabe, el valor de la varianza de un conjunto de datos no cambia si a todos éstos le sumamos (o restamos) un número constante. Esta "invarianza" de la varianza se debe a que no cambia su numerador, que es SC (la suma de los cuadrados de las desviaciones de los datos respecto a su media); en este caso el numerador de Q es SC_{grupos} . Su valor no cambia si restamos 1 a cada total de los grupos G_i que contienen exclusivamente 1's.
4. La eliminación de renglones (que es consecuencia de lo señalado en 2 y 3) hace que el cálculo de Q "a mano" se simplifique, teniéndose cuidado de que r no alcance el límite inferior de 4 que señalan empíricamente Tate y Brown (1970), de forma que la aproximación χ^2 cuadrado ya no sea buena. Esta disminución de renglones ayuda también en el tamaño del archivo de datos que lee el programa QCOCH para calcular Q .

Ejemplo

Remitiéndonos al ejemplo del uso de fuentes alimenticias por hormigas, presentado al inicio de esta sección, notamos que se desea comparar las $c = 4$ "fuentes alimenticias" respecto a los valores de una variable de respuesta dicotómica y_{ij} cuyos resultados se pueden codificar como "0" y "1" (0 = "no lo usa", 1 = "lo usa"). La matriz de datos

de las hormigas (Tabla 2) está formada por $r = 30$ renglones que son los bloques = "especies."

Las hipótesis nula y alternativa se escriben "coloquialmente" como:

H_0 : Las fuentes son usadas de igual manera por las especies.

H_A : Hay diferencia en el uso de las fuentes por parte de las especies.

"Matemáticamente", H_0 y H_A se interpretan haciendo $p_{ij} = P(y_{ij} = 1)$ (la probabilidad poblacional de que la fuente alimenticia j sea usada por la especie i). El enunciado "igual manera del uso de la fuente" que se señala en H_0 , significa que se cumple:

$$H_0: p_{i1} = p_{i2} = p_{i3} = p_{i4}, \text{ para cada } i=1, \dots, 30.$$

Es decir, que la probabilidad poblacional del uso de una fuente particular es igual a la probabilidad poblacional de uso de cualquier otra fuente, para cada especie, *pero puede variar de especie a especie* (de renglón a renglón).

La "diferencia en el uso de las fuentes" significa primero, que existe algún par de fuentes j y k y existe alguna especie i en donde las probabilidades poblacionales de uso de las fuentes son diferentes. Es decir:

$$H_A: p_{ij} \neq p_{ik}, \text{ para algún } j \text{ y } k \text{ y para algún } i.$$

Al correr el programa ejecutable QCOCH (el cual puede solicitarse al autor de estas notas), el archivo de datos se basó en 16 especies, al eliminar aquellas que no usaban fuente alguna o las que usaban todas las fuentes alimenticias (véanse las observaciones 2 y 3). Es así que el número "efectivo" de renglones (especies) fue $r = 16$. Se obtuvo $Q = 8.7736$, y un valor $P = 0.0320$ para la prueba de significación ji cuadrado con 3 grados de libertad.

5 Métodos y aplicaciones adicionales a la prueba Q de Cochran

1. Si sucede, como en el caso del ejemplo 2, que se rechaza la hipótesis nula de la prueba Q de Cochran, entonces el paso siguiente es realizar la prueba de comparación múltiple respecto al "número promedio de éxitos en cada grupo" para ver cuál de las fuentes difieren en cuanto a su uso (cuáles son la j y k que se señalan en la hipótesis alternativa). En Zar (1984, secc. 13.11), se describe el procedimiento propuesto por Marascuilo y McSweeney (1967) para la comparación múltiple de datos dicotómicos, que ilustramos a continuación con los datos del ejemplo 2.

Ejemplo 2 (Conclusión)

Consideremos las pruebas de comparaciones por parejas de medias de las $c = 4$ fuentes alimenticias. Para tal efecto, nos basamos en las fórmulas dadas por Zar (1984). Denotemos con $\gamma_j, j = 1, \dots, c (= 4)$ las medias poblacionales de tales grupos.

Sus estimaciones correspondientes son:

$$\bar{G}_j = \frac{G_j}{r}, \quad j = 1, \dots, c$$

Las hipótesis nulas que se prueban son:

$$H_0: \gamma_j - \gamma_k = 0, \quad j \neq k.$$

El estadístico de prueba para cada comparación es:

$$S_{jk} = \frac{|\bar{G}_j - \bar{G}_k|}{EE}$$

donde EE es el error estándar de la diferencia estimada de medias por grupo, dado por:

$$EE = \sqrt{2 \left(\frac{c \sum_j B_j - \sum_j B_j^2}{cr^2(c-1)} \right)}$$

Para un nivel de significación preestablecido α , sobre el conjunto completo de comparaciones, S_{jk} se compara con el valor crítico $\sqrt{\chi_{\alpha, c-1}^2}$. Si $S_{jk} \geq \sqrt{\chi_{\alpha, c-1}^2}$, se rechaza H_0 para las correspondientes γ_j y γ_k . Con los datos del ejemplo 2, usemos una tasa de error por experimento de $\alpha = 10$. Numerando las columnas que identifican a los grupos en la Tabla 2, del 1 al 4, tenemos:

$$r = 30, c = 4 \quad G_1 = 19 \quad G_2 = 21 \quad G_3 = 20 \quad G_4 = 13$$

$$\bar{G}_1 = 0.6333 \quad \bar{G}_2 = 0.7000 \quad \bar{G}_3 = 0.6667 \quad \bar{G}_4 = 0.4333$$

$$\sum_{i=1}^{30} B_i = 73, \quad \sum_{i=1}^{30} B_i^2 = 239$$

$$EE = \sqrt{2 \left(\frac{4(73) - 239}{4 \cdot 30^2 \cdot 3} \right)} = 0.009$$

Sustituyendo los valores correspondientes a S_{jk} , para las 6 parejas posibles, obtenemos:

$$S_{1,2} = 0.6734 \quad S_{1,3} = 0.3367 \quad S_{1,4} = 2.0202$$

$$S_{2,3} = 0.3367 \quad S_{2,4} = 2.6936$$

$$S_{3,4} = 2.3569$$

Como $\sqrt{\chi_{\alpha, c-1}^2} = \sqrt{\chi_{0.1, 3}^2} = \sqrt{6.251} = 2.5002$, el único par de medias de grupo que difiere significativamente es el (2,4). Es decir, tenemos evidencias de que las fuentes alimenticias "Néctar de otras estructuras reproductivas" y "Néctar extrafloral" son utilizadas de manera diferente por las especies de hormigas observadas ($P < 0.10$)

2. Es común encontrar situaciones en las que los resultados de un experimento o condiciones que se quieran comparar, se refieran a una variable dicotómica, una variable que es muy fácil de medir. Esta propiedad hace de la prueba Q de Cochran un método que encuentra aplicaciones diversas en Biología, Ciencias Sociales y Administración. Sin embargo, debe tomarse en cuenta que el estadístico Q se ha deducido suponiendo un modelo de permutación fundamental: que el total de éxitos

- en cada renglón es fijo y es igual a B_i . Si esta suposición cambia, obviamente la interpretación también.
3. Winer (1971) presenta la prueba Q de Cochran relacionando el problema de diseño en bloques para una variable dicotómica, con el de un diseño de mediciones repetidas y su análisis de varianza para datos expresados en forma de rangos (como la prueba de Friedman). En este contexto, la prueba Q de Cochran se basa en un modelo lineal aditivo de efectos mixtos: $y_{ij} = \mu + \tau_j + \beta_i + \varepsilon_{ij}$, en donde τ_j son efectos fijos de los grupos (tratamientos), con la restricción $\sum \tau_j = 0$, y los bloques forman, en conjunto, un efecto aleatorio, en donde $E(\beta_i) = 0$ y $\text{Var}(\beta_i) = \sigma_\beta^2$.
 4. En relación con las generalizaciones o modificaciones de la prueba Q de Cochran, podemos citar dos trabajos interesantes. El primero, de interés especial para los ecólogos, es un artículo de Hendrickson (1979), quien deduce el estadístico Q de una manera análoga a la presentada aquí (págs. 371-372). Además desarrolla un estadístico relacionado con la Q , que denota por M , basado en las coincidencias por parejas entre grupos. M también tiene una distribución asintótica simple (χ^2 con $(c-1)(c-2)/2$ grados de libertad.) Para motivar el cálculo de M , Hendrickson señala que en los estudios de comunidades, estudios de medios continuos (como ríos, pastizales, etc.) o estudios de monitoreo biológico, se acostumbra medir la asociación entre las especies (renglones de la matriz de 0's y 1's) halladas en los sitios (columnas) en términos de índices comparativos, como las coincidencias positivas por parejas de Jaccard. La interpretación práctica de un valor M significativo es que existe evidencia de que las especies en común en los sitios, no son independientes.
 5. La otra generalización que es digna de mención es la extensión que hizo Madansky (1963) de la prueba Q al caso en que las respuestas nominales tienen más de dos categorías (no son dicotómicas, sino "politómicas"). No obstante, su uso no es tan frecuente, posiblemente porque no se ha difundido lo suficiente.

Bibliografía

- BLOMQUIST, N. 1951. "Some tests based on dichotomization". *Annals of Mathematical Statistics*. 22, 362-271.
- COCHRAN, W.G. 1950. "The comparison of percentages in matched samples". *Biometrika*. 37:256-266.
- CONOVER, W.J. 1971. *Practical Nonparametric Statistics*. Wiley, New York. págs: 195-202.
- DMC Software Inc. 1992. ASP. Statistical Package for Business, Economics and the Social Sciences. Grand Blanc, Minnesota, U.S.A.
- HENDRICKSON, J.A. 1979. "Analyses of species occurrences in community, Continuum and biomonitoring studies". En: G.P. Patil y M. Rosenzweig, (eds.). *Contemporary Quantitative Ecology and Related Ecometrics*, pp. 361-397. *Statistical Ecology*, Vol. 12. International Co-Operative Publishing House, Maryland, USA.
- MADANSKY, A. 1963. "Test of homogeneity for correlated samples". *Journal of the American Statistical Association*. 58, 97-119.
- MARASCUULO, A y M. MCSWENEY, 1967. "Nonparametric post hoc comparisons for trend". *Psychological Bulletin*, 67 (Sept.), pp. 401-412.
- MCNEMAR, Q. 1947. "Note on the sampling error of the difference between correlated proportions or percentages". *Psychometrika*, 12 (June), pp. 153-157.
- MENDENHALL, W., R. SCHEAFFER, y D. WACKERLY. 1986. *Estadística Matemática con Aplicaciones*. Grupo Editorial Iberoamérica, México.
- PATIL, K. 1975. Cochran's Q Test: Exact Distribution. *Journal of the American Statistical Association*. 70, 349 (March), pp. 186-190.

- RICO-GRAY, V. 1993. "Use of Plant-Derived Food Resources by Ants in The Tropical Lowlands of Coastal Veracruz, México". *Biotropica* 25(3): 301-315.
- SINCICH, T. 1993. *Statistics by Example*. 5th. ed. Dellen-Macmillan, New York.
- TATE, M.W. y S.M. BROWN. 1970. "Note on the Cochran Q test". *J.Amer. Statist. Assoc.* 65:155-160.
- WALSH, J.E. 1947. "Concerning the effect of interclass correlation on certain significance tests". *Annals of Mathematical Statistics*. 18, 88.
- WINER, B. 1971. *Statistical Principles in Experimental Design*. McGraw-Hill, New York.
- ZAR, J.H. 1984. *Biostatistical Analysis*, 2d. ed. Prentice-Hall: Englewood Cliffs, N.J. U.S.A.
-