

PRUEBA EXACTA PARA TABLAS DE DOBLE ENTRADA CON MUESTRAS PEQUEÑAS Y PROPORCIONES EXTREMAS*

Dr. Jaime R. Robles

RESUMEN

Se introduce un procedimiento sugerido para tablas de doble entrada en psicología aplicada, incluyendo áreas clínicas y organizacionales. Se propone el uso de la prueba exacta de Fisher como una solución para situaciones con muestras pequeñas y proporciones extremas, típicas de variables como desórdenes psíquicos de baja prevalencia. El procedimiento es ilustrado con ejemplos tomados del área clínica y se propone una estrategia que envuelve el uso de la prueba exacta combinada con la razón de ventaja, discutiendo de manera introductoria las propiedades de esta estrategia de análisis.

Palabras Clave: Tablas 2X2, muestras pequeñas, proporciones extremas, prueba exacta de Fisher.

INTRODUCCIÓN

Es frecuente encontrar tratamientos de la prueba de Pearson para tablas de doble entrada, sin entrar en detalles como la aproximación de una distribución continua con la X^2 , a una distribución de tablas observadas, que es discreta en naturaleza. El automatismo, asistido por el análisis computarizado también contribuye a desestimar la importancia de discutir algunos fundamentos de la prueba de Pearson, a pesar de la gran variedad de opciones de pruebas disponibles (Oster, 2002).

Esto crea un escenario en el cual investigadores de distintos niveles de experticia y en diferentes áreas aplicadas de la psicología, se enfrentan a realidades concretas como

* Dedicado a la memoria de Luis Azagra L. s.j. y Andrés Miñarro Ll.

tablas observadas en muestras pequeñas, cuyas variables tienen proporciones extremas, ó lo que se denominan particiones sesgadas.

El propósito del presente artículo es presentar una introducción al análisis de este tipo de tablas de modo técnicamente apropiado, pero al mismo tiempo con relevancia para el investigador aplicado.

El tratamiento de temas complejos como la aproximación de una distribución continua a eventos discretos, la potencia de la prueba y métodos más complejos para pruebas exactas no condicionales, han sido mantenidos a un mínimo de complejidad para mantener el carácter introductorio del texto.

Por otra parte, las tablas usadas para ilustrar las nociones discutidas, son descritas en términos de frecuencias crudas, evitando el uso de porcentajes ó proporciones, para evitar las distorsiones introducidas por estos elementos en trabajos con muestras pequeñas. Para una introducción al uso de los porcentajes en las tablas de doble entrada, el lector es invitado a revisar textos como Kerlinger y Lee (2002).

En primer lugar se discuten algunos elementos del procedimiento convencional basado en la prueba de Pearson. Luego se discuten los elementos del procedimiento sugerido, con la prueba exacta de Fisher y el análisis de la razón de ventaja. Finalmente, se propone un procedimiento con pasos enumerados, para el tipo de tablas discutidas.

UBICUIDAD DE LA VARIABLE CATEGÓRICA EN ÁREAS APLICADAS

En áreas aplicadas es frecuente observar el uso de clasificaciones basadas en síntomas, indicadores y en general elementos diagnósticos discretos. En contraste con una aproximación estrictamente psicométrica en la cual se usan mediciones de rasgos ó continuos latentes, en muchos estudios vinculados a la psicología clínica, asesoramiento y áreas aplicadas en general, las categorías diagnósticas pueden considerarse omnipresentes.

Esto hace que exista un conjunto de variables categóricas al momento de realizar investigaciones en estas áreas. Un ejemplo sería la variable Depresión, la cual se puede medir como una puntuación en una escala ó se puede usar un conjunto de criterios y establecer una categorización diagnóstica binaria, de presente-ausente, como en el caso del sistema de códigos ICD ó DSM. Al usar estos sistemas categóricos, el análisis más frecuente al momento de la investigación es el de las tablas de doble entrada. (Agresti, 2001).

Las tablas de contingencia, en teoría pueden tomar cualquier forma desde 2×2 hasta $R \times C$, entendiendo que se pueden tener R filas y C columnas en la tabla, siendo R y C números relativamente grandes. No obstante, en muestras clínicas particularmente, es frecuente observar bajo número de casos y eventos de ocurrencia extrema.

Esta combinación hace que tener un gran número de sub-categorías diagnósticas y sub-clasificaciones en una misma tabla, resulte en tablas con muy bajo número de casos en cada celda ó cero casos en algunas celdas. En estos casos es mejor fundir las subcategorías y crear clasificaciones simples de 2 niveles, lo cual resulte en tablas 2×2 . Esto quiere decir que la tabla 2×2 es de especial utilidad en el común caso de muestras pequeñas y eventos de ocurrencia extrema (por ejemplo, trastornos de baja prevalencia).

PRUEBA DE PEARSON PARA TABLAS DE DOBLE ENTRADA.

Una idea de mucha trascendencia para comprender el análisis de las tablas de doble entrada es la de la tabla observada como un evento particular, dentro de un grupo de tablas posibles, dadas las variables cruzadas en la tabla. Una tabla de doble entrada observada no es un promedio de las posibles tablas ó ni siquiera un estimador de la tabla "verdadera", sino una tabla posible. Precisamente el análisis estadístico permite establecer cuán probable es observar esta tabla como resultado del azar. Para explicar mejor estas ideas, es conveniente introducir las nociones de los componentes de una tabla de doble entrada, y su análisis con la prueba convencional de Pearson.

En el resto del artículo se usarán tablas simuladas de acuerdo a un estudio nacional en los Estados Unidos (Kessler y cols, 2005), que muestra tasas de prevalencia *de corto plazo para Depresión Mayor (MDD) y Desorden de Pánico (PAN) son 6.7% y 2.7%. De acuerdo al mismo estudio y consistente con la literatura, hay alta comorbilidad entre desórdenes afectivos y de ansiedad.* En este sentido, suponiendo que del total de casos con MDD, 30% tiene comorbilidad con PAN, la tabla observada con una muestra de tamaño 50 sería la Tabla 1.

La tabla 1 muestra varias propiedades importantes. Llámese n_{11} al número de datos de la celda de la columna 1, fila 1, R_1 al total de la fila 1 y C_1 al total de la columna 1, mientras que N es el número total de observaciones. En este caso, $n_{11}=1$, $R_1=3$, $C_1=1$ y $N=50$.

Tabla 1: Depresión Mayor (MDD) y Desorden de Pánico (PAN) (N=50)

		PAN		Total
		Sí	No	
MDD	Sí	1	2	3
	No	0	47	47
	Total	1	49	50

La tabla tiene 1 grado de libertad y puede ser caracterizada por una celda, ya que al conocer el valor de esta celda y el de los marginales, el resto de las celdas quedan determinadas. Por ejemplo, dado que $n_{11}=1$ y $C_1=1$, necesariamente $n_{21}=0$, ya que el total de la columna debe sumar 1. Es fácil ver cómo los valores de las 3 celdas restantes quedan determinados al conocer los marginales (totales de fila y columna) y el valor de n_{11} . De esta forma, dados los marginales y N, la tabla puede ser identificada por el valor de n_{11} .

Dadas estas prevalencias y este tamaño muestral, existe un número limitado de tablas posibles. Para determinar cuántas tablas son posibles de observar, dados unos marginales, simplemente se determina el menor de los marginales, esto es $\min(R_1, R_2, C_1, C_2)$, y el número de tablas posibles corresponde a este valor +1, es decir el número de tablas posibles se comienza a contar desde 0. En la Tabla 1, el marginal menor es $C_1=1$, por lo cual existen sólo dos tablas a observar, $n_{11}=0$ (no casos comórbidos) y $n_{11}=1$, el ejemplo presentado. Esto ilustra cómo el número de tablas posibles, dados unos marginales, es limitado y que las distintas tablas posibles son entidades discretas.

En este sentido, el análisis de una tabla de doble entrada, tomando en cuenta los marginales, se denomina aproximación condicional (a los marginales observados). Existen otras aproximaciones semi-condicionales e incondicionales, usando estimadores poblacionales de los marginales, por ejemplo (Agresti y Min, 2002). No obstante, dichas aproximaciones implican una serie de supuestos estadísticos sumamente fuertes y no se consideran apropiadas en el contexto clínico ó aplicado discutido en el presente artículo.

Del ejemplo presentado anteriormente, se puede entender cómo las tablas de contingencia, incluyendo las 2x2 siguen una distribución discreta. Existe un número

limitado de tablas posibles, dados unos marginales y en este caso, cada tabla posible puede ser caracterizada por el valor de n_{11} .

Sin embargo, por conveniencia en el uso de los recursos de la inferencia estadística, usualmente la prueba estadística es hecha con referencia a la distribución teórica X^2 (Ji Cuadrado, o "Chi" Cuadrado), la cual es una distribución continua. Esto se hace con el propósito de aproximar la probabilidad de la tabla usando las propiedades convenientes de la distribución X^2 , ampliamente usada en esta y otras aplicaciones. De este modo, la prueba estadística de Pearson, se interpreta como distribuida como una variable X^2 , por lo que muchas veces se le denomina "Prueba de X^2 ".

En términos más precisos, la prueba de Pearson, es una prueba de independencia para tablas de doble entrada, cuya distribución teórica esperada es una distribución X^2 .

Para mostrar la inconveniencia de la prueba de Pearson para el análisis de la tabla 1, es necesario introducir el concepto de frecuencia esperada. Esta se determina por la siguiente fórmula

$$\text{Ecuación 1: } e_{ij} = (R_i \times C_j) / N$$

Siendo e_{ij} , la frecuencia esperada para la celda de la fila i y la columna j . De esta forma $e_{11} = (3 \times 1) / 50 = 0.06$. En la tabla 1, se tiene que el 75% de las celdas tienen $e_{ij} < 5$, y una de las celdas es 0, incumpliendo dos de los supuestos de la prueba estadística. Véase que estas propiedades tienen que ver mucho más con lo extremo de las particiones de las variables (en este caso las prevalencias) que con el tamaño muestral en sí mismo.

Para obtener el valor de la prueba estadística, se usa la siguiente fórmula

$$\text{Ecuación 2: } \chi = \sum [(n_{ij} - e_{ij})^2 / e_{ij}]$$

En el caso de una tabla 2×2 , asumiendo la Hipótesis nula de independencia (H_0), χ se distribuye teóricamente de acuerdo a una distribución X^2 , con 1 grado de libertad (Agresti, 1984). Esta prueba de Pearson, asume varias condiciones en las tablas, entre ellas, que la frecuencia esperada de todas las celdas sea mayor a 5 y que no existan celdas vacías (Cochran, 1954). En escenarios en los cuales se trabaja con proporciones de ocurrencia como las correspondientes a las prevalencias de MDD y PAN, es perfectamente esperable que resulten muchas tablas que no cumplan con estos supuestos.

Existen otras variaciones de esta prueba, como la prueba de razón de verosimilitud. Sin embargo, la aparente sencillez conceptual de la Ecuación 2 contribuye con su uso extendido, sin el debido razonamiento acerca de la aproximación a una distribución discreta, desde una distribución continua.

La tabla 1 es un ejemplo de cómo el uso de la aproximación por distribución continua a la distribución de las tablas con muestras pequeñas y variables sesgadas, resulta en incumplimiento de los supuestos de la prueba estadística, lo cual en sí es una razón para buscar una alternativa. Por otra parte, la aproximación por distribución continua tiene los problemas asociados a la interpolación de los valores de probabilidad y la prueba de Pearson, en sí misma, no provee de una métrica apropiada para evaluar la magnitud de la asociación.

INDEPENDENCIA Y MAGNITUD DE LA ASOCIACIÓN: LA RAZÓN DE VENTAJAS

Para entender mejor el valor de un análisis preciso de las tablas 2x2, es conveniente introducir una distinción entre la significancia de la prueba estadística y la magnitud de la asociación.

El análisis clásico de las tablas de doble entrada, pone a prueba la hipótesis estadística de independencia, que equivale a decir que la tabla observada es altamente probable en una extracción al azar. Estrictamente hablando, lo que se concluye de una prueba estadística significativa es la falta de independencia. No obstante, es frecuente encontrar el razonamiento que la falta de independencia implica asociación.

Adicionalmente, la falta de independencia (indicada por la significancia estadística), no es un indicador preciso de la magnitud de la asociación. Para estos efectos se puede usar la razón de ventajas (Odds Ratio en Inglés)

$$\text{Ecuación 3: } q = (n_{11} * n_{22}) / (n_{12} * n_{21})$$

La razón de ventajas tiene la propiedad de ser interpretable en términos crudos, por ejemplo, un valor de 2 significa que los casos con una condición 1 tienen 2 veces el chance de tener la otra condición (2), con respecto a los que no tienen la condición 1.

La tabla 2 muestra las mismas proporciones de la tabla 1, pero con un tamaño muestral de 150. Véase cómo, a pesar de tratarse de un tamaño muestral relativamente grande, aún la tabla tiene frecuencias bajas en varias de las celdas. Por otra parte, la

tabla 2 representa uno de 5 resultados posibles (desde 0 hasta 4 casos comórbidos), dados los marginales.

Tabla 2: Depresión Mayor (MDD) y Desorden de Pánico (PAN) (N=150)

		PAN		Total
		Sí	No	
MDD	Sí	3	7	10
	No	1	139	140
	Total	4	146	150

En la tabla 2, los casos con depresión tienen $q=59.57$ veces más riesgo de sufrir desorden de pánico, que los no depresivos. Esto puede ser mucho más informativo a nivel clínico ó aplicado, que la probabilidad de la tabla interpretada aisladamente.

Existen pruebas estadísticas e intervalos de confianza para q (Agresti, 1984), pero en general, usada en conjunto con la significancia de la prueba de Pearson ó la prueba de Fisher, se puede hacer un razonamiento adecuado.

La razón de ventajas q puede ser combinada con el análisis de los porcentajes direccionales (fila y columna), para hacer un análisis completo de la tabla observada. Por ejemplo, en la tabla 2, los casos comórbidos totalizan un 2%. No obstante, cuando se analizan los porcentajes de columna, puede verse que el 75% de los casos con PAN, tienen MDD, siendo consistente con el alto valor de q .

Cada celda en la tabla hace una contribución a q . Cuando una o más de las celdas tiene un valor desproporcionado, con respecto a los marginales, produce una desviación del valor de q con respecto a 1, indicando que hay una asociación de cierta magnitud. Esto puede ilustrarse claramente mediante el uso de un diagrama de cuadrantes ("Fourfold Plot") para representar las tablas de doble entrada (Friendly, 1994).

En la figuras 1a y 1b, se presentan los diagramas para la tabla 2 observada ($n_{11}=3$) y una tabla esperada por azar sin asociación ($n_{11}=0$). El diagrama presenta una serie de tres anillos concéntricos. El anillo central representa la contribución de cada celda a q (ó más precisamente a $\text{Log}(q)$), mientras que los anillos interior y exterior representan el intervalo de confianza.

Si los valores de las celdas son proporcionales, dados los marginales, el anillo central debe coincidir para los cuadrantes, dentro del margen de concordancia establecido por el intervalo de confianza. Este es el caso de la figura 1b. Cuando hay una asociación de cierta magnitud, se observa que el anillo central se descuadra, mostrando el valor desproporcional de las distintas celdas, como es el caso de la figura 1a.

Figura 1: Diagrama de cuadrantes (Fourfold Plot)

Figura 1a: Para la Tabla 2

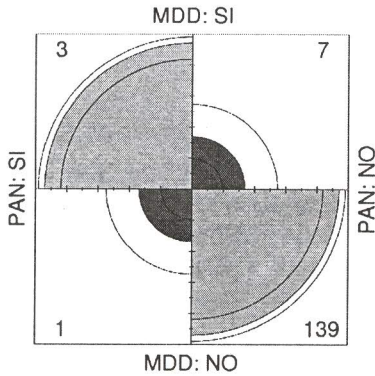
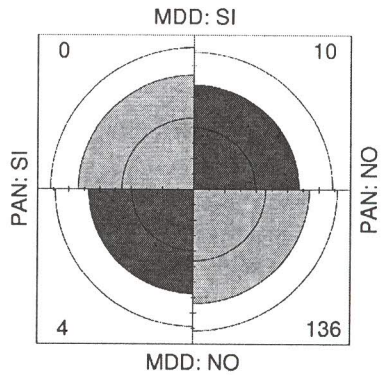


Figura 1b: Para una tabla sin asociación



En el caso de las particiones extremas, como las de la Tabla 2, es difícil evaluar los valores de las celdas, sin tomar en cuenta su proporcionalidad con respecto a los marginales. Esto queda claramente ilustrado en los diagramas de cuadrantes. Análisis similares pueden lograrse mediante el uso de porcentajes direccionales (Kerlinger y Lee, 2002).

PRUEBA EXACTA PARA TABLAS 2x2

La prueba exacta de Fisher, a veces llamada “probabilidad exacta de Fisher”, forma parte de una aproximación a la prueba estadística que incluye las pruebas exactas y las pruebas de permutación (Fisher, 1935).

Esta prueba puede considerarse como un procedimiento básico para la inferencia exacta, existiendo otros procedimientos mucho más complejos (Booth y Butler, 1999; Agresti, 2001). Sus usos son bastante extendidos, desde el convencional análisis categórico hasta incorporación en modelos para datos experimentales (Vargas-Irwin y cols, 2006).

Esta familia de pruebas exactas y de permutación, se ha hecho más accesible, dados los recursos computacionales en la actualidad. En algunos casos estas pruebas, y en particular la prueba de Fisher, son denominadas como pruebas no paramétricas ó pruebas de distribución libre. No obstante, esto puede ser impreciso. Hay una distribución teórica detrás de la prueba exacta, pero los supuestos que se hacen son mínimos, comparados con la prueba de Pearson. Por otra parte, los cálculos se hacen tomando estrictamente los datos observados, sin extrapolaciones a poblaciones ideales. En este sentido, la prueba exacta es más apropiada para muestras pequeñas, ó que pudieran pertenecer a subpoblaciones particulares, tales como muestras clínicas u organizacionales.

En el caso de las tablas de doble entrada, como las 2x2 tratadas aquí, la prueba exacta de Fisher se basa en una distribución discreta, que caracteriza la distribución de probabilidad de las tablas. Dicha distribución se denomina Distribución Hipergeométrica Extendida ó Hipergeométrica Extendida Multivariada (MXH, Plackett, 1981).

La función de probabilidad para la distribución MXH, de una tabla 2x2 t se define como:

$$\text{Ecuación 3: } p(t) = \frac{[R_1! R_2! (C_1! C_2!)]}{[N! (n_{11}! n_{12}! n_{21}! n_{22}!)]}$$

Tratándose de una distribución discreta, los cálculos involucran cálculo combinatorio (Boyett,1979; Patefield,1981; Saunders,1984), con uso extendido del operador factorial (!). Esto trajo problemas históricos en los computadores personales y lenguajes de programación comunes, por ejemplo, en un lenguaje convencional, con un procesador de 32 bits, el máximo número entero que se puede manejar es 232, limitando los valores que se pueden usar en esta fórmula.

Esto llevó al desarrollo de algoritmos de aproximación, como el algoritmo de redes (Pagano y Halvorsen, 1981; Metha y Patel, 1983), y el uso de programas especiales como StatXact, así como una gran variedad de algoritmos alternativos (Borkowf, 2004). No obstante, con procesadores y lenguajes de programación actuales han hecho posible que la prueba exacta esté disponible en programas más comunes como SPSS, SAS y R (Oster, 2002). Los cálculos presentados en este artículo son hechos con un software especial, que implementa un algoritmo para cálculos combinatorios que permite la solución de la prueba exacta aún para números muy grandes (Robles, 1996,2005).

El uso más frecuente de la prueba exacta es para inferencia bidireccional o de dos lados, lo cual es equivalente a contrastar H_0 , o lo que es lo mismo, hipótesis de $q=1$.

En ese caso, la probabilidad exacta asociada a la tabla, se computa sumando el valor de la función MXH para todas las tablas con valores iguales ó menores a la tabla en cuestión. En el caso de la tabla 2, la probabilidad bidireccional para la tabla con 3 casos comórbidos $p=0.000829+0.000010=0.000839$, dado que al valor de MXH de la tabla con 3 casos comórbidos (0.000829), hay que sumarle el valor MXH de la tabla con 4 casos comórbidos (0.000010).

En el caso de la tabla 2, la prueba arroja un valor de $p=0.000839$, significativo aún al nivel de significancia del 1/1000.

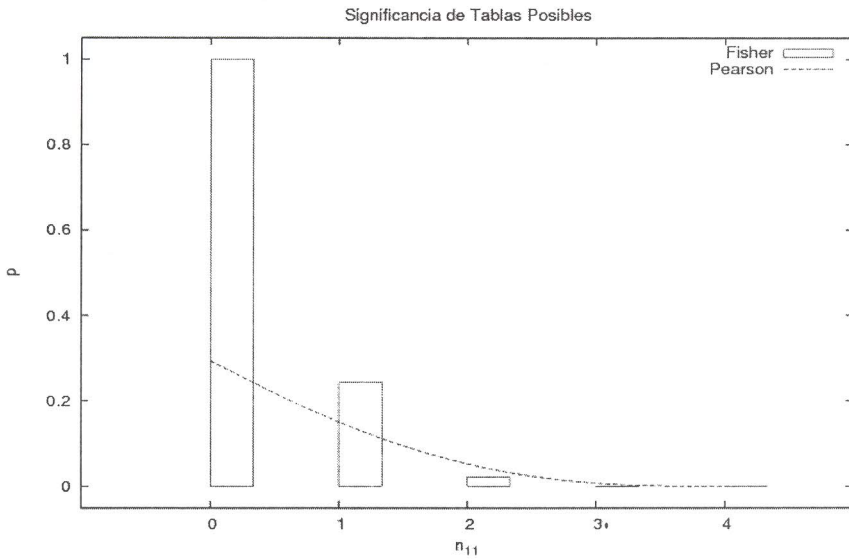
En la actualidad, el cómputo de la prueba exacta vía procedimiento manual es inusual, de cualquier modo queda ilustrado el sentido de la prueba de bidireccional, valor que hay que buscar en la salida de los programas como SAS y SPSS, que dependiendo de las opciones seleccionadas, pueden arrojar los valores para una y dos direcciones.

Este procedimiento es mucho más apropiado que la prueba de Pearson, ya que la tabla muestra e_{11} y $e_{22} < 5$, dado lo desequilibrado de las particiones, siendo las prevalencias de MDD y PAN valores bajos para crear una tabla susceptible de análisis por el procedimiento convencional de Pearson.

Por otra parte, como se discutió en la sección anterior, la magnitud de la asociación en este caso es notable, ya que $q=59.57$. Esto coincide con el valor extremo de la significancia arrojada por la prueba exacta. Esto se hace sin los inconvenientes de la prueba de Pearson y sin asumir supuestos estadísticos fuertes.

En este punto, se puede establecer una comparación entre los resultados de ambas pruebas (Pearson y Fisher). En la figura 2, se muestra el contraste entre las probabilidades discretas estimadas usando la prueba exacta de Fisher y la aproximación continua basada en la prueba de Pearson. En la distribución discreta de las tablas para este escenario, hay una tabla con probabilidad p_1 , con n_{11} casos comórbidos, mientras que la probabilidad de la tabla con $n_{11}+1$ casos comórbidos es p_2 . No hay tablas con valores intermedios entre p_1 y p_2 .

Figura 2: Probabilidades discretas estimadas usando la prueba exacta de Fisher y la aproximación continua basada en la prueba de Pearson.



En contraste, haciendo una aproximación a una distribución continua X^2 , existe un número teóricamente infinito de valores de probabilidad entre p_1 y p_2 . La tabla observada puede concebirse como un estimador poblacional, que no siempre es lo deseable, pero en todo caso, aun con muestras muy grandes, existe un pequeño número de tablas con probabilidades entre p_1 y p_2 , dadas las prevalencias y la comorbilidad. Esto hace que cuando se trabaja con la aproximación por distribución continua, se pueden hacer una serie de precisiones acerca de valores de probabilidad que no se traducen en diferencias prácticas, ya que mientras existe un inmenso rango de valores de probabilidad dentro de una precisión dada, sólo existe un mucho menor número de tablas posibles que caen dentro de ese rango de probabilidades.

El carácter condicional de la prueba, que toma en cuenta los valores de los marginales observados, puede ser problemático para la inferencia estadística, así como los saltos discretos entre probabilidades de las tablas. Esto ha llevado a algunos autores a desarrollar procedimientos especiales, como la interpolación de valores entre los puntos discretos (Metha y Walsh, 1992; Davison y Wang, 2002). No obstante, para efectos de investigación aplicada, el hecho de basarse en los datos observados puede ser considerado una ventaja, además del uso de supuestos mínimos a nivel estadístico.

La prueba exacta es calificada como una prueba conservadora, en el sentido que usualmente rechaza como no significativas asociaciones que son significativas de acuerdo a otros procedimientos (Agresti, 2001). Esto está estrechamente relacionado con el carácter discreto de MXH. Aparte del uso de interpolación de probabilidades, otra alternativa al uso de la prueba exacta, es el uso de algún procedimiento de simulación, para crear una distribución empírica de referencia, y así evitar el carácter conservador de la prueba exacta. No obstante, en un reciente estudio (Robles y van den Oord, 2006), se ha comprobado que el uso de la simulación simple, en lugar de la prueba exacta, puede resultar en falsos negativos, ó una estrategia aún más conservadora, dado que se descartarían asociaciones estadísticamente significativas de acuerdo con la prueba exacta, como no significativas de acuerdo con la distribución empírica derivada de la simulación.

PROCEDIMIENTO SUGERIDO

El caso general con muestras clínicas, estudios aplicados ó aún algunos estudios organizacionales, se tienen los mencionados componentes: Muestras pequeñas y particiones desequilibradas. Tomando en cuenta lo antes discutido, se sugiere el siguiente procedimiento:

1. Evaluación de q : Establecer la magnitud de la asociación a partir de la razón de ventajas.
2. Evaluar supuestos de la prueba de Pearson, si se cumplen, examinar la significancia de acuerdo a la distribución X^2 .
3. Si no puede usarse la prueba de Pearson, emplear la Prueba Exacta de Fisher (bidireccional). Esta prueba puede reemplazar la de Pearson en muestras pequeñas aún cuando se cumplan los supuestos de la prueba de Pearson.
4. En caso de asociación, interpretar valor de q acompañada de análisis de porcentajes direccionales.

COMENTARIOS FINALES

En las tablas de contingencia en contextos aplicados, es importante no hacer abstracción de la tabla observada, y usar un Análisis Razonado, tomando en cuenta elementos tales como riesgo relativo, comorbilidad, asociación, etc. La prueba estadística, especialmente la aproximación por distribución continua, es una abstracción para aproximarse a la probabilidad de la tabla observada, pero no substituye el análisis detallado de la tabla observada.

Otro aspecto de gran importancia, en el caso específico de una investigación en áreas clínicas es la atención a la magnitud de la asociación, ya que la significancia estadística en sí misma puede responder a preguntas de naturaleza más teórica. En las tablas 2x2, la magnitud de la asociación puede ser representada de varias maneras, con múltiples indicadores cuantitativos. No obstante, se puede usar q como indicador de magnitud de la asociación, sin necesidad de introducir nuevos indicadores. La métrica cruda de q es especialmente útil, dado que es muy claro su valor como indicador de asociación. Tomando en cuenta que la prueba exacta puede ser entendida como una prueba estadística de la hipótesis de $q=1$, el uso de la razón de ventaja combinado con la prueba exacta resulta en una estrategia consistente.

El procedimiento sugerido responde a la pregunta básica de la significancia estadística de la tabla observada y a una evaluación inicial de la magnitud de la asociación. Para análisis más complejos ó tablas multidimensionales, se recomienda el uso de modelos Log-lineales ó modelos logísticos (Agresti, 1984; Kerlinger y Lee, 2002).

El poder computacional disponible en la actualidad, puede llevar al automatismo de realizar la prueba de Pearson sin mayores consideraciones. En lugar de esto, debe aprovecharse el poder de cómputo para examinar alternativas, como la prueba exacta, que en el pasado eran prohibitivas por su complejidad de cómputo.

REFERENCIAS

- Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. New York: John Wiley & Sons.
- Agresti A. (2001). Exact inference for categorical data recent advances and continuing controversies. *Statistics in Medicine*, 20:2709-2722.
- Agresti A, Min Y. (2002). Unconditional small-sample confidence intervals for the odds ratio. *Biostatistics* 3:379-386.
- Borkowf , C.B. (2004). An efficient algorithm for generating two-way contingency tables with fixed marginal totals and arbitrary mean proportions, with applications to permutation tests. *Computational Statistics & Data Analysis* 44:431-449.
- Boyett, J.M. (1979). Algorithm AS144. Random $R \times C$ tables with given row and column totals. *Applied Statistics* 28: 329-332.

- Booth, J.G. ; Butler, R.W. (1999). An importance sampling algorithm for exact conditional tests in log-linear models. *Biometrika* 86: 321-332.
- Cochran, W.G. (1954). Some methods of strengthening the common chi-square tests. *Biometrics* , 10: 417-451.
- Davison, A.C. ; Wang, S. (2002). Saddlepoint approximations as smoothers. *Biometrika*, 89: 933-938.
- Fisher, R. A. (1935). The logic of inductive inference. *Journal of the Royal Statistical Society A*, 39-54.
- Friendly, M. (1998). Conceptual Models for Visualizing Contingency Table Data. En: Balsius, J. y Greenace, M. *Visualization of Categorical Data*. San Diego: Academic Press.
- Kerlinger, F. N.; Lee, H. B. (2002). *Investigación del comportamiento. Métodos de Investigación*. México: McGraw-Hill.
- Kessler, R.C.; Chiu, W.T.; Demler, O.; Walters, E.E. (2005) Prevalence, Severity, and Comorbidity of 12-Month DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62, 617-627.
- Metha, C.R.; Patel, N.R. (1983). A network algorithm for performing the Fisher exact test in $r \times c$ contingency tables. *Journal of the American Statistical Association*, 78: 427-434.
- Metha, C.R. ; Walsh, S.J. (1992). Comparison of exact, mid-p, and Mantel-Haenszel confidence intervals for the common odds ratio across several 2×2 contingency tables. *American Statistician*, 46: 146-150.
- North, B.V.; Curtis, D.; Sham, P.C. (2002). A note on the calculation of empirical p values from Monte Carlo procedures. *American Journal of Human Genetics*, 71: 439-441.
- Oster, R.A. (2002). An examination of statistical software packages for categorical data analysis using exact methods. *American Statistician*; 56:235-246.
- Pagano, M.; Halvorsen, K.T. (1981). An algorithm for finding the exact significance levels of $r \times c$ tables. *Journal of the American Statistical Association*, 76: 931-934.

- Patefield, W.M. (1981). Algorithm AS159. An efficient method of generating random $R \times C$ tables with given row and column totals. *Applied Statistics*, 30: 91-97.
- Plackett, R.L. (1981). *The Analysis of Categorical Data*, 2nd Edition. New York, NY: Macmillan Publishing Co., Inc.
- Robles, J.R. (1996). PRS Polytomous Item Generation-Simulation according to the common-factor model. *Applied Psychological Measurement*, 20: 140.
- Robles, J.R. (2005). A Java program to compute probabilities according to the Multivariate Extended Hypergeometric Distribution. Virginia Institute of Psychiatric and Behavioral Genetics, Richmond, Virginia.
- Robles, J.R.; van den Oord, E. (2006). A cautionary note on the use of simulation procedures for analyzing contingency tables containing a small expected cell frequency, *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 141B:414-417.
- Saunders I.W. (1984). Algorithm AS205. Enumeration of $R \times C$ tables with repeated row totals. *Applied Statistics* 33:340-352.
- Vargas-Irwin, C., Beardsley, P. M., van den Oord, E.J.C.G. and Robles, J. R. (2006). A Method for Analyzing Strain Differences in Acquisition of IV Cocaine Self-administration in Mice. *Behavior Genetics*, 36, 525-535.

APENDICE: Ejemplo de Código en SPSS

* Código para Tabla 2 en SPSS

* Probado con SPSS 14.x y superior

* Debe funcionar con otras versiones con cambios mínimos

* Código de ejemplo que se distribuye sin ninguna garantía

**

* Primero se crea la tabla de ejemplo con los datos agrupados

* La tercera columna de datos contiene las frecuencias.

DATA LIST LIST /MDD(F8.0) PAN(F8.0) FREC(F8.0).

BEGIN DATA

1 1 3

1 0 1

0 1 7

0 0 139

END DATA. .

LIST.

* Muy importante: ponderar la tabla por las frecuencias.

WEIGHT BY FREC.

*Para la prueba de Pearson solamente usar \STATISTICS=CHISQ

*En su lugar, se usa la opción ALL para obtener el Odds-Ratio

*En este caso SPSS computa Fisher automáticamente

CROSSTABS VARIABLES=MDD(0,1) PAN (0,1)

 /TABLES=MDD BY PAN

 /CELLS=TOTAL EXPECTED

 / STATISTICS=ALL.

*incluir /CELLS=ROW COL para obtener porcentajes direccionales.