

COMPRENSIÓN DE ASOCIACIÓN EN TABLAS DE CONTINGENCIA POR ESTUDIANTES DE PSICOLOGÍA

Gustavo R. Cañadas, Carmen Batanero, María M. Gea, José M. Contreras
Facultad de Educación, Universidad de Granada

Aceptado 1 junio 2013. Revisado 28 junio 2013. Publicado 2 julio 2013

RESUMEN

Las tablas de contingencia son un instrumento importante de resumen de la información, aprendizaje de conceptos y análisis de datos, no obstante diversas investigaciones describen las dificultades de los estudiantes. En este trabajo, se describe una experiencia de enseñanza de las tablas de contingencia, evaluando el aprendizaje de uno de los temas en una muestra de 93 estudiantes del primer año del grado de Psicología. Se presentan a modo de evaluación los resultados en 6 ítems de opción múltiple y un problema abierto, que indican un aprendizaje satisfactorio, adquisición de estrategias formales y superación de la concepción local y causal sobre la asociación, concepciones descritas por Estepa (1994). La evaluación tenía una finalidad diagnóstica y los resultados se usarían para proponer acciones educativas adecuadas que mejorarían esta enseñanza en el futuro. Se realiza una descripción de las estrategias y soluciones aportadas por los estudiantes en las preguntas realizadas en el problema abierto.

Palabras clave: Tablas de contingencia, dificultades, materiales, formación psicológica

UNDERSTANDING OF ASSOCIATION IN CONTINGENCY TABLES FOR PSYCHOLOGY STUDENTS

Contingency tables are an important tool for summarizing the information, learning and data analysis concepts; however several studies described the difficulties of the students. In this paper, we describe a teaching experience of contingency tables, assessing learning a topic in a sample of 93 students the first year of Psychology degree. Is presented by way of assessment results in six multiple-choice items and an open problem, which indicate successful learning, acquisition of formal strategies and overcoming local causal conception about the association, conceptions described by Estepa (1994). The evaluation had a diagnostic purpose and the results would be used to propose appropriate educational activities to improve the teaching in the future. We present a description of the strategies and solutions provided by students in the questions asked in the open problem.

Key words: Contingency tables, difficulties, teaching materials, psychological training

Las tablas de contingencia se presentan con frecuencia en el diagnóstico, evaluación psicológica, y otras tareas profesionales, y sirven para evaluar si existe asociación entre las variables correspondientes (Díaz y Gallego, 2006). Sin embargo, la investigación psicológica y didáctica muestra frecuentes sesgos en su interpretación, aunque, a pesar de ello se presta poca atención a este tema en la formación de psicólogos.

Las tablas de contingencia tienen un papel esencial como instrumento de transnumeración, una forma básica de razonamiento estadístico que proporciona nueva información, al cambiar de un sistema de representación a otro (Wild y Pfannkuch, 1999). Sin embargo se presta poca atención a este tema en la educación universitaria, en el supuesto de que su comprensión es sencilla.

En este trabajo se describe una experiencia de enseñanza con materiales desarrollados en nuestro grupo de investigación; se analiza también la evaluación del aprendizaje de uno de los temas. En lo que sigue se comienza analizando la investigación previa y se describe el método y resultados.

Investigaciones previas

Inhelder y Piaget (1955) inician la investigación sobre la asociación, describiendo las estrategias en los juicios de asociación de los chicos a partir de 13 años. Para explicar sus resultados utilizan una tabla similar a la Tabla 1, indicando que en una primera etapa, sólo se usa la celda a (presencia de los caracteres A y B) sin comprender que la celda d (ausencia de los dos caracteres) tiene el mismo peso en relación a la asociación. En una segunda etapa (entre los 13 y 15 años), se comparan las celdas dos a dos (por ejemplo a con b, o a con c). El tercer paso, sería comprender cuales son los casos favorables (a y d) y desfavorables (b y c) de la asociación, sin

compararlos y finalmente, el sujeto establecen las relaciones diagonales, comparándolas entre sí o con el total (a+b+c+d).

Posteriormente autores, como Smedlund (1963) y Shaklee y Mins (1982) han analizado las estrategias de sujetos adultos, concluyendo que, en contraposición con lo supuesto por Piaget e Inhelder, algunos continúan con estrategias propias de niños, por lo que el desarrollo del concepto de asociación no sería espontáneo en los sujetos. La dificultad de la tarea la muestran Jenkins y Ward (1965), indicando que incluso la estrategia de comparar las diagonales, considerada como correcta por Inhelder y Piaget, sólo llevaría a un juicio de asociación correcto, si las frecuencias marginales por filas o por columnas en la Tabla 1 fuesen iguales, pero no en el caso general. La estrategia correcta para el caso general consiste en comparar la diferencia entre las probabilidades $P(B|A)$ y $P(B|no A)$.

Tabla 1. Tabla de contingencia 2x2

	A	No A	Total
B	a	b	a+b
No B	c	d	c+d
Total	a+c	b+d	

Pérez Echeverría (1990) clasifica dichas estrategias en niveles de dificultad (niveles 1 a 3 si se usan sólo 1 a 3 de las 4 celdas de la tabla; nivel 4 si se usan las 4 celdas con estrategias aditivas y nivel 5 si se usan las 4 celdas con estrategias multiplicativas).

Otro problema es que muchas personas forman sus propias teorías sobre la relación entre variables en la tabla de contingencia que les impide evaluar correctamente la asociación (Chapman y Chapman, 1969), fenómeno conocido como "correlación ilusoria".

Por otro lado, Estepa (1994) analiza las concepciones erróneas de los estudiantes y describe la concepción causal, según la cual el sujeto sólo considera la asociación entre variables si puede adjudicarse a la presencia de una relación causal entre las mismas. También define la concepción unidireccional donde el estudiante no admite la asociación inversa y la local, cuando el sujeto basa su juicio en sólo una parte de los datos.

Experiencia de enseñanza

Con la finalidad de contribuir a superar los anteriores problemas, se desarrollaron una serie de didácticos para facilitar la comprensión de las tablas de contingencia en estudiantes universitarios de ciencias sociales. Estos materiales fueron experimentados en el curso 2010-2011 (Cañadas, 2011) dentro de dos grupos de clase regulares de primer curso de Psicología (en total 94 alumnos). El tiempo dedicado a la enseñanza fue 6 sesiones de 1 hora de duración. Cuatro de dichas sesiones se llevarían a cabo en el aula tradicional en los grupos mencionados y se dedicarían a la presentación de los temas. También se dedicaron dos sesiones prácticas en el laboratorio de informática en las cuáles cada alumno trabaja independientemente con el ordenador utilizando unas hojas de Microsoft Excel, preparadas para realizar las actividades prácticas. Cada grupo de teoría se dividió en tres subgrupos de aproximadamente 15 alumnos, todos ellos impartidos por el mismo profesor. El contenido se ha organizado en cuatro lecciones:

1. *Tablas de contingencia, lectura e interpretación.* Se trata que los alumnos aprendiesen a: (a) Resumir datos sobre dos variables estadística en una tabla de contingencia; (b) Identificar las frecuencias dobles (c) Calcular las frecuencias relativas dobles, marginales y condicionales e interpretarlas; (d) Representar gráficamente los datos mediante diagrama de barras adosadas, diagrama de barras apiladas y gráfico tridimensional y (e) Calcular probabilidades simples, compuestas y condicionales a partir de datos de una tabla de contingencia.
2. *Asociación estadística, dependencia funcional e independencia.* Se trató de que los alumnos aprendiesen a: (a) Diferenciar la asociación estadística, dependencia funcional e independencia; (b) Reconocer el tipo de relación entre dos variables comparando las frecuencias condicionales; (c) Calcular las frecuencias esperadas en caso de independencia y (d) Analizar posibles explicaciones de una asociación estadística: relación causal, interdependencia, tercera variable explicativa o asociación espuria.
3. *El estadístico Chi-cuadrado y contrastes asociados.* La tercera lección, se dedicó a: (a) Dar una medida de la diferencia entre frecuencias observadas y esperadas en caso de independencia; (b) Calcular e interpretar el estadístico Chi-cuadrado y sus grados de libertad; (c) Comprender los pasos para llevar a cabo el contraste de independencia y el contraste

de homogeneidad y (e) Comprender los supuestos del contraste Chi-cuadrado.

4. *Medidas de asociación.* Se trató de que los alumnos aprendiesen a: (a) Interpretar la intensidad de la dependencia entre dos variables en una tabla de contingencia; (b) Calcular e interpretar medidas de asociación en tablas 2x2: Coeficiente Phi de Pearson, Riesgo relativo y Razón de productos cruzados; (c) Calcular e interpretar medidas de asociación en tablas rxc; Coeficiente de contingencia de Pearson y V de Cramer y (d) Calcular e interpretar medidas de reducción del error de predicción de una variable, cuando se conoce el valor de la otra: Lambda de Goodman y Kruskal.

Para asegurar la validez del estudio las clases fueron observadas. Las interacciones en la clase también fueron grabadas en audio, para posteriormente poder comparar con la observación y anotar las principales incidencias y dudas planteadas por los estudiantes. Los profesores habituales de los cursos también asistieron a las sesiones.

Evaluación

También se prepararon diferentes pruebas de evaluación, para ser utilizadas en tres momentos del proceso de estudio, en cuya construcción se han seguido las recomendaciones psicométricas habituales para asegurar su calidad. El aprendizaje se evaluó con un cuestionario, compuesto de 6 ítems para cada uno de los temas y algunos problemas abiertos. En este trabajo presentamos los resultados obtenidos en relación con el tema 2 "asociación estadística, dependencia funcional y aleatoria". Su comprensión fue evaluada con los 6 ítems que se reproducen a continuación, marcando en negrita las respuestas correctas en cada ítem y un problema abierto.

Resultados de la evaluación en ítems cerrados

El ítem 1 evalúa el conocimiento que el estudiante adquiere de propiedades sencillas que permiten evaluar si dos variables de una tabla de contingencia son independientes: Las opciones a) y b) son correctas, pues en caso de independencia, todas las distribuciones condicionales por fila o columna coinciden. La c) es igualmente correcta, ya que la independencia implica esta relación de igualdad entre las frecuencias condicionales y las frecuencias marginales.

Ítem 1. Para que dos variables de una tabla de contingencia sean independientes, han de ser iguales:

- a. Las frecuencias relativas condicionales por columnas.
- b. Las frecuencias relativas condicionales por filas.
- c. Las frecuencias relativas condicionales y frecuencias relativas marginales.

El ítem 2 evalúa el conocimiento del procedimiento de cálculo de las frecuencias esperadas en una tabla de contingencia. La respuesta correcta es

$$e_{ij} = \frac{f_{i.} \times f_{.j}}{n}$$

la c), pues $e_{ij} = \frac{f_{i.} \times f_{.j}}{n}$. En la respuesta a) se confunde las frecuencias absolutas dobles con las frecuencias absolutas marginales, en el cálculo de las frecuencias esperadas; en la b) se cambia las frecuencias relativas dobles por las frecuencias absolutas marginales, en el proceso de cálculo y en la d) se confunde las frecuencias relativas marginales con las frecuencias absolutas marginales, además de que falta el total de la muestra, en el cálculo.

Item 2. Las frecuencias esperadas se calculan mediante:

- a. Las frecuencias absolutas dobles y el total de la muestra.
- b. Las frecuencias relativas dobles y el total de la muestra.
- c. Las frecuencias absolutas marginales y el total de la muestra.**
- d. Las frecuencias relativas marginales

Item 3. En caso de que haya asociación entre variables, las frecuencias relativas dobles:

- a. En todas las celdas son iguales al producto del total por fila y columna que le corresponda, es decir $h_{i,j} = h_{i.} \cdot h_{.j}$
- b. Puede ocurrir que coincida en alguna celda al producto del total por fila y columna que le corresponda, es decir $h_{i,j} = h_{i.} \cdot h_{.j}$**
- c. Nunca son iguales al producto del total por fila y columna que le corresponda, es decir nunca se cumple $h_{i,j} = h_{i.} \cdot h_{.j}$

Evalúa el conocimiento de la relación entre frecuencias relativas y marginales en caso de independencia. La respuesta correcta es la b), pues en caso de asociación podría darse la igualdad en un caso. La respuesta a) es incorrecta, ya que describe una propiedad de la independencia y no de la asociación. Asimismo lo es la respuesta c), puesto que puede ocurrir que en caso de asociación en unos casos se cumpla esta igualdad y en otros no.

En el ítem 4 se desea conocer la interpretación que dan los alumnos a la relación existente entre diferentes celdas en la tabla 2x2 y el signo de la asociación. La respuesta adecuada es la b), pues en la diagonal principal están los valores que informan de dependencia directa, referidos a la presencia-presencia (A-B), y ausencia-ausencia (no A-no B). Por el contrario en las otras dos celdas, se da un solo carácter y el otro no y serían las celdas favorables a una asociación inversa. Las otras opciones son incorrectas.

Item 4. En las siguientes tablas 2x2 indicamos el tipo de asociación que informan las diferentes celdas ¿Cuales de las siguientes tablas es correcta?

a.

	B	No B
A	Dep. directa	Dep. directa
No A	Dep. directa	Dep. inversa

b.

	B	No B
A	Dep. directa	Dep. inversa
No A	Dep. inversa	Dep. directa

c.

	B	No B
A	Dep. directa	Dep. inversa
No A	Dep. inversa	Dep. inversa

Item 5. Indica cuál de las siguientes frases es cierta:

- a. Si hay una relación causal entre A y B, entonces habrá asociación positiva entre A y B
- b. Si al tomar datos de A y B encontramos asociación entre las variables, entonces habrá una relación causal entre A y B
- c. Si hay una relación causal entre A y B, entonces habrá asociación, que puede ser positiva o negativa.**

Este ítem trata de detectar la concepción causal de la asociación (Estepa, 1994). La respuesta adecuada es la c), pues si A es causa de B, entonces habrá asociación. La respuesta a) es incorrecta, pues causa si implica asociación, pero puede ser positiva o negativa. Asimismo es errónea la respuesta b); aquí aparece la falacia "correlación implica causalidad".

Item 6. La diferencia entre la dependencia funcional y la dependencia aleatoria consiste en:

- a. En la dependencia aleatoria a cada valor de la variable independiente X le corresponde sólo un valor de la variable dependiente Y
- b. En la dependencia aleatoria, al variar X suele variar Y, pero no siempre**
- c. La dependencia aleatoria puede ser directa o inversa, pero la funcional siempre es directa

Este ítem analizar la comprensión de la diferencia entre dependencia funcional y aleatoria. La respuesta correcta es la b), pues la dependencia aleatoria no implica un único valor de la variable dependiente al variar la independiente. La respuesta a) es incorrecta, pues da la definición de dependencia funcional. La respuesta c) es asimismo incorrecta, puesto que la dependencia funcional no es siempre directa.

En la Tabla 2 se resumen las respuestas a los ítems correspondientes al Tema 2. Estos resultados muestran un alto grado de aprendizaje, que se refleja en las pocas respuestas en blanco y el alto porcentaje de respuestas correctas. Los mejores resultados se obtienen en el ítem 4 (identificación de celdas que informan del signo de la asociación en la tabla 2x2); ítem 5 (diferencia entre asociación y causalidad) y opción a) del ítem 6 (dependencia funcional). Por otro lado, los mayores errores se producen en la opción b) del ítem 3 (no se reconoce que en caso de independencia en alguna celda puede ser igual las frecuencias esperadas y observadas y la opción c) del ítem 1 (confusión de una propiedad de independencia).

Tabla 2. Frecuencias (y porcentajes) de respuestas en los ítems (n=94)

	Apartado	Correcto	Incorrecto	En blanco
Ítem 1	a (Verdadera)	67 (71,3)	24 (25,5)	3 (3,2)
	b (Verdadera)	69 (73,4)	22 (23,4)	3 (3,2)
	c (Verdadera)	27 (28,7)	64 (68,1)	3 (3,2)
Ítem 2	a (Falsa)	74 (78,7)	18 (19,1)	2 (2,1)
	b (Falsa)	80 (85,1)	12 (12,8)	2 (2,1)
	c (Verdadera)	59 (62,8)	33 (35,1)	2 (2,1)
	d (Falsa)	87 (92,6)	5 (5,3)	2 (2,1)
Ítem 3	a (Falsa)	39 (41,5)	51 (54,3)	4 (4,3)
	b (Verdadera)	17 (18,1)	73 (77,7)	4 (4,3)
	c (Falsa)	59 (62,8)	31 (33)	4 (4,3)
Ítem 4	a (Falsa)	91 (96,8)	1 (1,1)	2 (2,1)
	b (Verdadera)	90 (95,7)	2 (2,1)	2 (2,1)
	c (Falsa)	91 (96,8)	1 (1,1)	2 (2,1)
Ítem 5	a (Falsa)	89 (94,7)	5 (5,3)	0 (0)
	b (Falsa)	80 (85,1)	13 (13,8)	1 (1,1)
	c (Verdadera)	79 (84)	15 (16)	0 (0)
Ítem 6	a (Falsa)	77 (81,9)	12 (12,8)	5 (5,3)
	b (Verdadera)	43 (45,7)	46 (48,9)	5 (5,3)
	c (Falsa)	60 (63,8)	29 (30,9)	(5,3)

Resultados de la evaluación en un problema abierto

También se propuso sobre este tema el siguiente problema abierto, que trata de evaluar la competencia del estudiante para establecer un juicio de asociación y sus estrategias

Problema. Un grupo de 200 personas aquejadas de ansiedad fue dividido aleatoriamente en dos subgrupos. Al primer grupo se ofreció un medicamento realmente efectivo y al otro se ofreció un placebo (medicamento sin efecto). Al cabo de un mes fueron interrogados sobre la eficacia conseguida, con el siguiente resultado:

	Su ansiedad ha disminuido	Siguen con mucha ansiedad
Medicamento	50	15
Placebo	96	39

- ¿Hay asociación entre el tipo de tratamiento (medicamento o placebo) y el efecto producido (la ansiedad disminuye o no)? ¿O son las variables independientes?
- Indica cómo has llegado a esta conclusión (puedes usar el método que prefieras)

En este problema, se plantean preguntas sobre una tabla 2x2, donde no existe una asociación entre las variables, ya que no es estadísticamente significativa ($p=0,386$). El alumno puede tener teorías

previas al respecto, pensando que el medicamento mejora la ansiedad y la relación sería de tipo interdependencia. Este problema fue resuelto sin ayuda del ordenador.

Solución prevista para el estudio de la asociación.

El estudiante debe hacer un juicio de asociación, a partir de los datos en una tabla de contingencia 2x2 (Batanero y Díaz, 2008). El estudiante ha de leer la tabla, identificar las variables del estudio: tipo de medicamento (contra la ansiedad o placebo) y si hay o no mejora.

El estudiante podría realizar un contraste formal de independencia, para lo cual ha de recordar que en un contraste de independencia se trata de decidir si las dos variables son ó no independientes. Tendría que plantear las hipótesis, calcular el estadístico Chi-cuadrado, que toma el valor 0,75, especificando sus grados de libertad que son: $(2-1) \times (2-1)=1$. Comparando con el valor crítico para 1 g.l., llegaría a la conclusión de que los resultados no son estadísticamente significativos, por lo cual habría que aceptar provisionalmente la hipótesis de independencia.

También podría usar cualquier procedimiento, incluso intuitivo. Por ejemplo, podría realizar una comparación entre las frecuencias esperadas en caso de independencia (Batanero y Díaz, 2008) con las frecuencias observadas, concluyendo que son muy parecidas; esto indica que no hay relación entre las variables; las pequeñas diferencias no son importantes. Asimismo el estudiante puede no hacer ningún tipo de cálculo, sino comparar los valores de las celdas que informan de una dependencia directa (50 y 39), con los valores de las celdas que informan de una dependencia inversa (15 y 96), según Inhelder y Piaget (1955), llegando a la misma conclusión que en casos anteriores. A continuación se describen las respuestas encontradas.

Estrategias utilizadas en el estudio de la asociación*Estrategias correctas:*

EC.1. Comparar las distribuciones condicionales por filas ó por columnas. Es decir, el estudiante aplica una de las propiedades de la independencia, la cual dice que si todas las distribuciones condicionales, por filas o por columnas, coinciden, entonces las variables son independientes. En consecuencia, el estudiante está usando todas las celdas de la tabla y razonando el nivel 5 de Pérez Echeverría (1990), pues emplea estrategias multiplicativas. Además hace un uso explícito de los conceptos de frecuencias condicionales e independencia. También fue descrita en el estudio de Estepa (1994) Un ejemplo sería el siguiente:

"He calculado las frecuencias relativas condicionales por filas y los resultados no han sido iguales, por lo tanto no hay independencia de las variables habiendo así asociación entre ellas" (Estudiante 82).

EC.2. El estudiante calcula las frecuencias esperadas en caso de independencia y las compara con las observadas. Si dichas frecuencias coinciden en todas las celdas, entonces considera que las variables son independientes. El estudiante utiliza explícitamente una de las propiedades de la independencia, así como los conceptos de frecuencias esperadas e independencia, lo que nos indica la emergencia de una estrategia formal, fruto del aprendizaje, pues no se ha descrito en los estudios

sobre juicios de asociación, como el de Estepa (1994). Un ejemplo aparece a continuación, donde resaltamos el hecho que, al no disponer el estudiante de ordenador, debe recordar la fórmula de cálculo de las frecuencias esperadas y realizar dicho cálculo con la calculadora:

"Frecuencias esperadas"

47,45	17,55
98,55	3,645

"Son independientes no existe relación entre estas variables" (Estudiante 10).

EC.3. El estudiante calcula el estadístico Chi-cuadrado para realizar un contraste de independencia, completando los pasos del proceso y tomando una decisión sobre la dependencia o independencia, según los resultados. Al igual que en las dos anteriores, esta sería una estrategia formal, fruto del aprendizaje. Además es una estrategia muy elaborada, en la que utiliza los conceptos de frecuencia observada y esperada, contraste de hipótesis, estadístico Chi-cuadrado y su distribución, nivel de significación, región de aceptación y rechazo y la lógica del contraste de hipótesis. A continuación presentamos una respuesta en esta categoría.

"Frecuencias esperadas"

47,45	17,55
98,55	3,645

0,14	0,37
0,12	0,18

$\chi^2=0,14+0,37+0,12+0,18=0,81$; no es estadísticamente significativo, por lo tanto son independientes" (Estudiante 93).

Estrategias parcialmente correctas:

EP.1. Comparación de razones de casos a favor y en contra. El estudiante el cual compara las

razones $\frac{50}{96}$ y $\frac{15}{39}$, de casos en los que la ansiedad ha disminuido en los que toman el medicamento y placebo y aquellos en que no ha disminuido, ya que en caso de independencia estas razones coinciden.

Es decir, aunque el estudiante no lo indica explícitamente, está basándose en uno de los coeficientes de asociación (razón de productos cruzados), que no llega a calcular e interpretar explícitamente, por lo que hemos considerado la estrategia parcialmente correcta. Sería una estrategia descrita por Inhelder y Piaget (1955) como propia de las operaciones formales, pues usa todas las celdas con comparaciones multiplicativas y de nivel 5 según Pérez Echeverría (1990), que nosotros también hemos reforzado en el aprendizaje. Un caso de este tipo es:

$$\frac{50}{96} \neq \frac{15}{39}$$

"96 \neq 39" (Estudiante 77).

Estrategias incorrectas:

El.1. El estudiante intenta aplicar un procedimiento informal, en el cual se comparan las celdas que informan de una dependencia directa, con las celdas que informan de una dependencia inversa, y al observar que la mayoría de los sujetos pertenece a uno de estos dos grupos toma una decisión. El conflicto aparece porque no usa todas las celdas de la tabla. Generalmente se usan dos celdas (presencia-presencia y ausencia-ausencia) que informarían que la asociación en la tabla es directa, según Inhelder y Piaget (1955). Esta es una estrategia de nivel 2 en la clasificación de Pérez Echeverría (1990). Un ejemplo es como sigue:

"Hay mayoría personas que cumplen las dos condiciones o no cumplen ninguna" (Estudiante 72).

El.2. Se calculan las frecuencias relativas dobles y se comparan entre sí. El estudiante quiere aplicar una de las propiedades de la independencia, donde si las frecuencias esperadas condicionales coinciden con las frecuencias absolutas, entonces hay independencia. Pero en lugar de utilizar las frecuencias esperadas utiliza las frecuencias relativas. Calcula la probabilidad de los sucesos A y B y trata de comparar con las probabilidades condicionales, confundiendo $P(B|A)$ con $P(A|B)$, error descrito por Falk (1996).

El estudiante tiene un conflicto en que confunde estas dos frecuencias, que también aparece en Díaz (2007) y Contreras (2011). Aunque es de nivel 5 (se usan todas las celdas y comparaciones multiplicativas) es incorrecta. Un ejemplo se incluye a continuación.

0,25	0,075
0,48	0,195

"Porque las frecuencias por filas o por columnas no dan los mismos resultados, por tanto son dependientes" (Estudiante 6).

El.3. El estudiante compara frecuencias marginales creyendo en forma errónea, que si las frecuencias marginales coinciden, entonces hay independencia. Esta estrategia también fue encontrada en Estepa (1994) y supone falta de comprensión de la necesidad de trabajar con frecuencias relativas al comparar distribuciones. Un ejemplo de este tipo de respuesta sería:

"Son dependientes, ya que ninguno de los resultados totales coincide" (Estudiante 43).

RI.4. Dividir las cuatro frecuencias absolutas por una frecuencia absoluta marginal. El estudiante quiere aplicar una de las propiedades de la independencia, donde si las distribuciones condicionales coinciden, entonces hay independencia. Pero el estudiante tiene un conflicto al calcular las frecuencias condicionales. Por ejemplo:

$$\begin{aligned} 50/65 &= 0,76 \\ 96/65 &= 0,7 \\ 15/65 &= 0,23 \\ 39/65 &= 0,6 \end{aligned}$$

0,76	0,23	0,28
0,7	0,28	0,675

"No son independientes. Sí hay asociación" (Estudiante 81).

El problema planteado presenta una asociación moderada y los estudiantes en su mayoría lo han reconocido (72 estudiantes, 76,5%). La mayoría también indica que la dependencia es directa (60 estudiantes, es decir 64,5%), mostrando de nuevo un buen aprendizaje. En la Tabla 3 se presentan las frecuencias de las estrategias utilizadas, la mayoría correctas; un 21% utiliza estrategias formales fruto del aprendizaje, un 17,1% estrategias intuitivas, pero todas ellas de nivel 5 en la clasificación de Pérez Echeverría; un 44,1% no indica la estrategia, posiblemente debido a falta de capacidad de argumentación, pues como se ha indicado la mayoría de estudiantes responde correctamente a la asociación y su signo.

Tabla 3. Frecuencias y porcentajes de respuestas

	Frecuencia	Porcentaje
RC.1	66	70,2
RC.2	4	4,2
RC.3	2	2,1
RP.1	3	3,2
RI.1	1	1,1
RI.1	1	1,1
RI.2	2	2,1
RI.3	1	1,1
No responde	14	14,8
Total	94	100

CONCLUSIONES

En este trabajo hemos descrito una experiencia de enseñanza de la asociación en tablas de contingencia, presentando un breve resumen de la evaluación del aprendizaje en uno de los temas. Los resultados indican un aprendizaje satisfactorio, y competencia en la detección de asociación en un problema abierto, con uso tanto de estrategias formales, como de estrategias de nivel 5 en la clasificación de Pérez Echeverría. Se superan las concepciones causal y local, descritas por Estepa (1994). Ha sido menor la capacidad de argumentación de los estudiantes, muchos de los cuáles no fue capaz de explicar su estrategia en el problema abierto. Como consecuencia, aunque el material se ha relevado de interés al usarlo con alumnos de Psicología, se plantea, así el reto de continuar este trabajo con nuevas investigaciones sobre la enseñanza y aprendizaje de las tablas de contingencia.

REFERENCIAS

- Batanero, C. y Díaz, C. (2008). *Análisis de datos con Statgraphics*. Granada: Departamento de Didáctica de la Matemática.
- Cañadas, G. R. (2011). *Las tablas de contingencia para Psicología*. Departamento de Didáctica de la Matemática, Universidad de Granada.
- Chapman, L. J. & Chapman, J.P. (1969). Illusory correlation as an obstacle to the use of valid Psychodiagnostic signs. *Journal of Abnormal Psychology*, 74, 271-280.
- Contreras, J. M. (2011). *Evaluación de conocimientos y recursos didácticos en la formación de profesores sobre probabilidad condicional*. Tesis Doctoral. Universidad de Granada.
- Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología*. Tesis doctoral. Universidad de Granada.
- Díaz, J. y Gallego, B. (2006). Algunas medidas de utilidad en el diagnóstico, *Revista Cubana de Medicina General Integrada*, 22(1).
- Estepa, A. (1994). *Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores*. Tesis Doctoral. Universidad de Granada.
- Inhelder, B. & Piaget, J. (1955). *De la logique de l'enfant à la logique de l'adolescent*. Paris: Presses Universitaires de France.
- Jenkins, H. M. y Ward, W. C. (1965). Judgment of the contingency between responses and outcomes, *Psychological Monographs*, 79, 1-17.
- Pérez Echeverría, M. P. (1990). *Psicología del razonamiento probabilístico*. Madrid: Ediciones de la Universidad Autónoma de Madrid.
- Shaklee, H. y Mins, M. (1982). Sources of error in judging event covariations: Effects of memory demands, *Journal of Experimental Psychology Learning, Memory and Cognition*, 8(3), 208-224.
- Smedlund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, 4, 165-174.
- Wild, C., y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (con discusión). *International Statistical Review*, 67(3), 223-265.