

COMPRESIÓN DEL TEST CHI-CUADRADO POR ESTUDIANTES DE PSICOLOGÍA

UNDERTANDING OF CHI-SQUARED TEST BY PSYCHOLOGY STUDENTS

Cañadas, G.R. ⁽¹⁾, **Batanero, C.** ⁽¹⁾, **Díaz, C.** ⁽²⁾, **Gea, M.M.** ⁽¹⁾

Universidad de Granada ⁽¹⁾, *Universidad de Huelva* ⁽²⁾

Resumen

En este trabajo describimos un estudio de las dificultades en la realización del contraste Chi-cuadrado de homogeneidad en estudiantes de psicología, después de un curso de análisis de datos. A partir de un problema abierto, resuelto con ayuda de un programa de cálculo Excel, se analiza el establecimiento de las hipótesis, identificación del estadístico y p-valor, decisión tomada e interpretación de resultados.

Abstract

The Chi- squared test is frequently used in psychology and education. This paper describes a study of the competence to carry out this test in psychology students after studying the subject. Using an open-ended problem that is solved with the help of an Excel program, we analyse the setting of hypotheses, identification of the statistics and p-value, decision taken and interpretation of results.

Palabras clave: *Contraste de hipótesis, Chi-cuadrado, comprensión.*

Key words: *Statistical tests, chi-squared, understanding.*

Introducción

Un análisis estadístico habitual es el contraste Chi-cuadrado, para estudiar la homogeneidad de muestras o la asociación en tablas de contingencia. Aunque la literatura previa ha analizado las dificultades de los estudiantes en el contraste de hipótesis (e.g., Vallecillos, 1994), dichos estudios se han basado en ítems de opción múltiple, o contrastes paramétricos sobre la media de una población normal y han evaluado la comprensión de conceptos, más que la competencia de los estudiantes para completar todos los pasos en este procedimiento.

Para completar la investigación, el objetivo de este trabajo es evaluar la forma en que los estudiantes de Psicología, después de haber estudiado el tema, llevan a cabo los pasos necesarios en un contraste Chi-cuadrado de homogeneidad: planteamiento de hipótesis, identificación del estadístico Chi-cuadrado, cálculo del p -valor utilizando software, decisión tomada e interpretación de resultados. Es parte de una investigación más amplia en la que se estudia el aprendizaje de las tablas de contingencia. En lo que sigue se expone el fundamento y método del trabajo y se discuten los resultados.

Fundamentos

Antecedentes

La mayoría de las investigaciones sobre contraste de hipótesis describen errores en la interpretación del nivel de significación y el p -valor (ej., Vallecillos, 1994), denunciándose el intercambio de los términos de la probabilidad condicional en su definición.

Respecto al concepto de hipótesis, Vallecillos (1994) y Batanero (2000) indican que la hipótesis nula se supone cierta, se plantea para ser rechazada y determina la distribución muestral del estadístico; pero los estudiantes asignan estas propiedades a la hipótesis alternativas o plantean sus hipótesis utilizando el estadístico, en lugar de usar el parámetro. Castro Sotos et al. (2007) y Vera, Díaz y Batanero (2011) informan sobre la dificultad en identificar la población bajo estudio, y en discriminar contrastes unilaterales y bilaterales. Vallecillos (1994) también indica que los estudiantes confunden el criterio de decisión al aplicar el contraste y, por tanto, llegar a una decisión equivocada. Esto relaciona con la dificultad de comprensión de la lógica subyacente al contraste de hipótesis (Harradine, Batanero y Rossman, 2011).

En las investigaciones anteriores, el contraste estudiado es el de la media muestral; no hemos encontrado estudios previos sobre el contraste de homogeneidad, que es el problema que se aborda en este estudio.

Marco teórico

El trabajo se apoya en el enfoque Onto-semiótico (Godino, Batanero y Font, 2007), más concretamente, en los tipos de objetos matemáticos propugnados en esta teoría: (a) Situaciones problemas (en nuestro trabajo decidir sobre la homogeneidad de una serie de muestras); (b) lenguaje (expresiones, símbolos y gráficos que se usan en el trabajo matemático, como H_0 y H_1 , usados para plantear las hipótesis); (c) Conceptos: por ejemplo, estadístico y parámetro, nivel de significación; (d) propiedades: por ejemplo, que las hipótesis nula y alternativa son complementarias; (e) procedimientos;

como el cálculo del valor Chi-cuadrado; y (d) argumentos: usados para justificar la solución del problema.

También utilizamos la idea de *conflicto semiótico*. En las prácticas matemáticas se requiere un uso continuo del lenguaje matemático, pues los objetos matemáticos son inmateriales; por ello a cada expresión matemática (objeto inicial, o signo) le corresponde un contenido (objeto final, esto es, lo representado; generalmente mediante un criterio o regla de correspondencia. El carácter relacional de la actividad matemática puede explicar algunas dificultades y errores de los estudiantes. Godino, Batanero y Font (2007) denominan conflicto semiótico a las interpretaciones de expresiones matemáticas por parte de los estudiantes que no concuerdan con las que el profesor trata de transmitir. Dichos conflictos semióticos producen equivocaciones en los estudiantes, que no son debidos a falta de conocimiento, sino a una interpretación incorrecta de expresiones matemáticas.

En este trabajo utilizaremos el método de análisis semiótico propuesto por estos autores, para clasificar las respuestas de los estudiantes en la solución del problema planteado. Este análisis consiste en la identificación las prácticas matemáticas al tratar de resolver el problema y de los objetos y procesos matemáticos implicados. Como resultado se identificarán algunos conflictos semióticos, que explican los errores de los estudiantes.

Método

La muestra estuvo formada por 92 alumnos que siguieron un curso de “Técnicas de análisis en la investigación psicológica” (primer año de psicología) el curso 2010-2011. Se dedicaron 4 sesiones teóricas y 2 prácticas a las tablas de contingencia, incluyendo el test Chi-cuadrado de homogeneidad. El problema propuesto (Figura 1) se completó como parte de la evaluación. Los estudiantes disponían de un programa de cálculo en Excel en el que, introduciendo la tabla, obtenían el valor Chi-cuadrado, aunque debían proporcionar los grados de libertad para obtener el p -valor. El problema plantea un contraste de homogeneidad. En el apartado (a) el estudiante ha de interpretar el enunciado, recordar que en un contraste de homogeneidad se trata de decidir si los datos provienen de diferentes poblaciones. Una posible formulación correcta de las hipótesis sería la siguiente:

H_0 : *Las tres palabras tienen idéntica distribución de componente emocional;*

H_1 : *La distribución de componente emocional de alguna palabra es diferente.*

En un experimento un psicólogo selecciona tres palabras y decide evaluar su componente emocional en una muestra aleatoria. Presenta, independientemente, cada palabra a 100 sujetos y registra si la palabra es percibida como emocionalmente positiva, negativa o neutra. A la vista de los resultados, ¿Podrá el psicólogo considerar que las tres palabras tienen el mismo componente emocional?

Componente emocional de la palabra	Palabras		
	1	2	3
Positivo	26	45	32
Negativo	32	27	38
Neutro	42	28	30

- Formule las hipótesis adecuadas para efectuar un contraste Chi-cuadrado de homogeneidad
- Calcule el estadístico de contraste, y la probabilidad asociada al valor obtenido
- Decida si se debe rechazar o no la hipótesis nula para $\alpha = 0,01$
- Interprete el resultado obtenido en el contexto de la investigación.

Figura 1. Problema planteado a los estudiantes)

Para resolver el segundo apartado, el estudiante dispone de un programa de cálculo, que proporciona el valor $\chi^2=10,81$, pero debe especificar los grados de libertad (4 g.l.). Una vez proporcionados, el programa calcula el p -valor= $0,029$. En el apartado (c) el alumno ha de recordar la lógica del contraste, la forma en que se construyen la región de aceptación y rechazo, la definición de valor p y nivel de significación y deducir que el resultado no es estadísticamente significativo para el valor crítico del problema. El último paso es interpretar los resultados: a nivel $0,01$ no podemos decir que alguna palabra tenga diferente componente emocional y en consecuencia, debemos aceptar que tienen el mismo componente.

Resultados

Recogidos los datos se realizó un análisis semiótico de las respuestas en cada apartado, encontrando las siguientes categorías de respuestas. Se denota como RCn las respuestas correctas, RPn las parcialmente correctas y RIn las incorrectas.

Formulación de hipótesis

En la Tabla 1 se presenta un ejemplo de cada categoría, que son las siguientes:

RC.1. El alumno usa una notación adecuada para la hipótesis nula y alternativa, y hace referencia explícita al parámetro “proporción poblacional” en forma simbólica al establecer las hipótesis, elegidas en forma correcta.

RC.2. Hace referencia al concepto de distribución de la variable (en vez del parámetro poblacional) y no usa expresión simbólica. Las hipótesis nula y alternativa están correctamente formuladas.

Tabla 1. Ejemplos de cada categoría del apartado a

Categoría	Ejemplo
RC.1	$H_0: p_{1j} = p_{2j} = p_{3j}$ H_1 : Alguna subpoblación es diferente
RC.2	H_0 : Todas las subpoblaciones tienen idéntica distribución para la variable Y . H_1 : La distribución de la variable Y en alguna de estas subpoblaciones es diferente
RP.1	$H_0: \mu_1 = \mu_2 = \mu_3$ $H_1: \mu_n \neq \mu_m$, para algún n, m perteneciente $\{1, 2, 3\}$
RP.2	H_0 : Las tres palabras tienen el mismo componente emocional H_1 : Las palabras tienen distintos componentes emocionales
RP.3	H_0 : Componente emocional1 = Componente emocional2 = Componente emocional3 $H_1: CE_1 \neq CE_2 \neq CE_3$
RI.1	H_0 : Las palabras presentadas influyen en las emociones H_1 : Las palabras presentadas no influyen en las emociones
RI.2	H_0 : 10,81
RI.3	H_0 : Las variables están asociadas. H_1 : Las variables son independientes
RI.4	$H_0: X_1 \neq X_2 \neq X_3$ $H_1: X_1 = X_2 = X_3$
RI.5	H_0 : Las variables entre columnas no están asociadas H_1 : Las variables entre filas si están asociadas
RI.6	$H_0: W_1 = W_2 = W_3$ (W =palabra). $H_1: W_1 \neq W_2 \neq W_3$
RI.7	$H_0: p_{1j} = p_{2j}$ H_1 : son distintos
RI.8	$H_0: p_{1j} = p_{2j} = p_{3j}$ $H_1: p_{1j} \neq p_{2j} = p_{3j}$

RP.1. Se elige correctamente la hipótesis nula (como de igualdad de poblaciones) y la alternativa (como de diferencia entre alguna de ellas); pero se hace referencia a la media poblacional, en vez de usar la proporción. Esto último es incorrecto, pues no tiene sentido usar la media en datos cualitativos.

RP.2. Lenguaje poco preciso donde no se puede deducir si el alumno distingue la variable de su distribución., aunque formula correctamente la hipótesis nula. No hace referencia al concepto de población o distribución muestral, que según Harradine, Batanero y Rossman (2011) son más abstractos que los de población o muestra.

RP.3. Elige correctamente la hipótesis nula, pero la alternativa no es correcta, ya que sólo se exige la diferencia entre alguna de ellas; Al igual que en Vera, Díaz y Batanero (2011), las dos hipótesis no cubren el espacio paramétrico.

RI.1. Confunde las hipótesis nula y alternativa (error descrito por Vallecillos, 1994 y Chow, 1996) y expresa el contraste como de independencia, error razonable, debido a que ambos contrastes se realizan a partir de tablas de contingencia y usando el estadístico Chi-cuadrado.

RI.2. Presenta el estadístico Chi-cuadrado como hipótesis nula, sin dar la hipótesis alternativa. Confunde parámetro y estadístico, puesto que plantea la hipótesis en término de un estadístico, error descrito por Vallecillos (1994) y Castro Sotos et al. (2007).

RI.3. Confunde el contraste de homogeneidad con el de independencia. También intercambia hipótesis nula y alternativa, error señalado en Vallecillos (1994) y Chow (1996).

RI.4. Intercambia hipótesis alternativa y nula. Además en la hipótesis alternativa indica que las tres poblaciones han de ser diferentes, por lo que las hipótesis no cubren el espacio paramétrico (Vera, Díaz y Batanero, 2011).

RI.5. Confunde el contraste de homogeneidad con el de independencia e indica que la asociación se da entre filas o en columnas. De hecho, confunde valor y variable, error señalado por White (1980).

RI.6. Confunde la variable bajo estudio (componente emocional) con las poblaciones (datos tomados sobre las palabras). Además las dos hipótesis no cubren el espacio paramétrico (Vera, Díaz y Batanero, 2011).

RI.7. Usa dos poblaciones (en vez de tres); posiblemente confunde el número de variables con el de las subpoblaciones que se comparan, observándose dificultad para identificar la población bajo estudio (White, 1980).

RI.8. Plantea bien la hipótesis nula, pero en la alternativa sólo considera que sean distintas las dos primeras distribuciones, por lo que las dos hipótesis no cubren el espacio paramétrico (Vera, Díaz y Batanero, 2011).

Tabla 2. Resultados en el apartado a

Respuesta		Frecuencia	Porcentaje
No responde		7	7,6
Correcta	RC.1	29	31,4
	RC.2	14	15,2
Parcialmente correcta	RP.1	10	10,9
	RP.2	7	7,6
	RP.3	12	13,1
Incorrecta	RI.1	1	1,1
	RI.2	1	1,1
	RI.3	1	1,1
	RI.4	4	4,3
	RI.5	2	2,2
	RI.6	1	1,1
	RI.7	2	2,2
	RI.8	1	1,1
Total		92	100

En la Tabla 2 observamos 41,4% plantean correctamente las hipótesis, 26,8% en forma parcialmente correcta, y 13,7% incorrectamente. Las respuestas más frecuentes son las correctas (RC.1 y RC.2), en la que plantean las hipótesis de un contraste de homogeneidad, una de forma simbólica y otra verbalmente; y las parcialmente correctas

(RP.1), donde hace referencia a la media muestral, en vez de usar la media poblacional, y (RP.3) que en la hipótesis alternativa es incorrecta. Los resultados son mejores que los de Vallecillos (1994), donde Vera, Díaz y Batanero donde 26% y 33% de los participantes plantean correctamente las hipótesis.

Estadístico de contraste y probabilidad asociada

En la Tabla 3 se presentan ejemplos de cada tipo de respuesta que se analizan a continuación.

Tabla 3. Ejemplos en cada categoría del apartado b

Categoría	Ejemplo
RC.1	Chi-cuadrado experimental = 10,81; p -valor = 0,029
RP.1	Chi-cuadrado experimental = 10,81; p -valor = 0,021
RI.1	Chi-cuadrado experimental = 10,81; p = 0,001
RI.2	Chi-cuadrado experimental = 10,81; p = 13,28
RI.3	Chi-cuadrado experimental = 10,81
RI.4	Chi-cuadrado experimental = 10,81; p = 0,01

RC.1 El alumno usa una notación adecuada, obtiene el estadístico Chi-cuadrado, los grados de libertad y el valor p correcto.

RI.1 El alumno calcula correctamente el Chi-cuadrado pero confunde los grados de libertad y obtiene por tanto un valor p incorrecto. Esta confusión de los grados de libertad también se encuentra en las investigaciones de Alvarado (2007) y Olivo (2008), en el primer caso en el trabajo con la distribución t y en el segundo, con las distribuciones, t , Chi-cuadrado y F .

RI.2 El alumno calcula correctamente el valor Chi cuadrado y los grados de libertad. Obtiene en la tabla de la distribución el valor crítico 13,28, confundiendo con el p -valor, sin notar que una probabilidad no puede ser mayor que la unidad, error descrito por Contreras et al. (2010).

RI.3 El alumno obtiene el Chi-cuadrado pero no reconoce, el valor de la probabilidad asociada al valor obtenido, es decir de p .

RI.4 El alumno da el alfa, en vez de calcular el valor p , confundiendo estos dos conceptos, error descrito por Vallecillos (1994).

Tabla 4. Resultados en el apartado b

Respuesta		Frecuencia	Porcentaje
Correcta	RC.1	59	64,2
Incorrecta	RI.1	14	15,2
	RI.2	6	6,5
	RI.3	12	13
	RI.4	1	1,1
Total		92	100

Los resultados se presentan en la Tabla 4; la mayoría de los estudiantes obtienen a partir del programa correctamente el estadístico y la probabilidad asociada (55,8%).

Resaltan entre las respuestas incorrectas (RI.1), en la que no recuerdan los grados de libertad y (RI.3), donde el alumno no reconoce el valor p error descrito por Alvarado (2007) y Olivo (2008), en el uso de la distribución t , Chi-cuadrado y F .

Toma de decisión

En este apartado sólo hemos encontrado dos tipos de respuestas; correcta e incorrecta. Son respuestas correctas aquellas en que el estudiante interpreta adecuadamente los resultados, y no rechaza la hipótesis nula con formulaciones como: “*no podemos rechazar la homogeneidad*”. Es errónea la de los alumnos que deciden rechazar la hipótesis, confundiendo los criterios de aceptación y rechazo. En la Tabla 5 observamos que la mayoría (51,9%) toman la decisión correcta, siendo mejores los resultados que en los de investigaciones de Vallecillos (1994) aunque casi la mitad de los alumnos no llegan a una decisión correcta, error también señalado por Haller y Kraus (2002).

Tabla 5. Respuestas en el apartado c

	Frecuencia	Porcentaje
Correcta	54	51,9
Incorrecta	30	32,6
No responde	8	8,7
Total	92	100

Interpretación de resultados

A continuación se describen las categorías de respuestas (ver ejemplos en la Tabla 6).

Tabla 6. Ejemplos en cada categoría del apartado d

Categoría	Ejemplo
RC.1	El componente emocional se distribuye homogéneamente en las diferentes palabras
RP.1	Al buscar en la tabla de Chi-cuadrado obtenemos un valor de $p=0,029$. Por tanto no es estadísticamente significativo pues es mayor que 0,01. Aceptamos la hipótesis de homogeneidad
RI.1	Rechazamos H_0 , esto significa que las tres palabras tienen diferente componente emocional.
RI.2	Rechazaríamos la hipótesis de independencia (porque este valor tan cercano a 0 indica un valor muy improbable de Chi-cuadrado)
RI.3	Como el valor de $p=0,029$ es muy cercano a 0, indica un valor muy improbable de Chi-cuadrado si la H_0 es cierta, rechazamos la igualdad de componente emocional
RI.4	Observando el resultado del estadístico Chi cuadrado obtenemos que la probabilidad de obtener un valor 13,28 o mayor con 4 grados de libertad es $p=0,029$

RC.1. Los alumnos interpretan los resultados en el contexto tomando una decisión correcta.

RP.1. Los alumnos no interpreta en el contexto del problema como se les pide, pero toman una decisión correcta haciendo uso de una interpretación genérica de contraste de hipótesis.

RI.1. Los alumnos interpretan en el contexto del problema, pero la interpretación que hacen es incorrecta, por tomar una decisión incorrecta, pues rechazan la hipótesis nula.

RI.2. Los alumnos no interpretan en el contexto del problema, y además toman una decisión incorrecta de la toma de decisión.

RI.3. Los alumnos interpretan en el contexto del problema, pero toma una decisión incorrecta. Además estos alumnos tienen en cuenta el valor p , pero no α explícitamente.

RI.4. Estos alumnos no interpretan ni en el contexto, ni de ninguna forma. Además, no toman una decisión.

Tabla 7. Resultados en el apartado d

	Frecuencia	Porcentaje
RC.1	40	43,5
RP.1	5	5,4
RI.1	20	21,7
RI.2	2	2,2
RI.3	1	1,1
RI.4	1	1,1
No responde	23	25
Total	92	100

En la interpretación de los resultados según el contexto del problema (Tabla 7), observamos un 43,5% que lo hacen en forma correcta, siendo la respuesta más frecuente. El resto de estudiantes no llega a interpretar el resultado en relación al contexto.

Discusión y conclusiones

En el problema planteado los estudiantes recorren todos pasos del contraste de hipótesis; 79,3% plantean hipótesis correctas o parcialmente correctas, 64,1% determinan correctamente el estadístico y valor p (único punto facilitado por el software), 51,9% toman la decisión correcta y 43,5% interpretan correctamente los resultados en contexto, recorriendo, por tanto todos los pasos del proceso de modelización (Henry, 1997): planteamiento de hipótesis, definición y trabajo con un modelo matemático e interpretación de resultados.

Los alumnos fueron capaces de abstraer y simplificar la realidad, pasando de la idea de componente emocional, a la distribución de la percepción de los 100 participantes, esto es, aceptando los datos del problema como indicadores de dicha componente emocional. Trabajaron con modelos matemáticos (las variables aleatorias y estadísticas correspondientes a la percepción del componente emocional, el estadístico Chi-cuadrado, sus grados de libertad, valor p y nivel de significación, lógica del contraste de hipótesis, etc.). Finalmente el último paso (interpretar los resultados del

trabajo matemático realizado con el modelo (decisión tomada) en el contexto del problema (traducir esta decisión a lo que indica respecto de la componente emocional de las palabras) fue lo más difícil, de acuerdo a lo obtenido en Arteaga, Batanero y Ruiz (2009).

En consecuencia se plantea la necesidad de continuar la investigación respecto a la enseñanza del contraste de hipótesis y la modelización estadística. Sería necesario igualmente insistir en los aspectos interpretativos en la enseñanza, pues los profesores se concentran con frecuencia en el trabajo matemático (y no en las fases iniciales y finales de la modelización) que son igualmente importantes.

Agradecimientos

Proyecto EDU2010-14947 y becas FPU-AP2009-2807 y BES-2011-044684 (MCINN-FEDER) (MCINN- FEDER) y grupo FQM126 (Junta de Andalucía).

Referencias

- Alvarado, H. (2007). *Significados del teorema central del límite en la enseñanza de la estadística en ingeniería*. Tesis Doctoral. Universidad de Granada, España.
- Arteaga, P., Batanero, C. y Ruiz, B. (2009). Comparación de distribuciones por futuros profesores. En M. J. González, M.T. González y J. Murillo, *Investigación en Educación Matemática XIII* (pp. 129-138). Santander: SEIEM.
- Batanero, C. (2000). Controversies around the role of statistical tests in experimental research. *Mathematical Thinking and Learning*, 2 (1-2), 75-98.
- Castro Sotos, A. E., Vanhoof, S., Van den Nororgate, W., Onghena, P. (2007). Student's misconceptions of statistical inference: A review of the empirical evidence from research on statistical education. *Educational Research Review*, 2 (2), 98-113.
- Chow, L. S. (1996). *Statistical significance: Rational, validity and utility*. London: Sage.
- Contreras, J. M., Estrada, A., Díaz, C. y Batanero, C. (2010). Dificultades de futuros profesores en la lectura y cálculo de probabilidades en tablas de doble entrada. En M. Moreno, A. Estrada, J. Carrillo y T. Sierra (Eds.). *Investigación en Educación Matemática XIV* (p. 271-280). Lleida: SEIEM.
- Godino, J. D. Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education*, 39 (1-2), 127-135.
- Haller, H. y Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research*, 7(1). On line: www.mpr-online.de/.
- Harradine, A., Batanero, C. y Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education*. Springer
- Henry, M. (1997). Notion de modèle et modélisation en l'enseignement. *En Enseigner les probabilités au lycée* (pp. 77-84). Reims: Commission Inter-IREM.
- Olivo, E. (2008). *Significados de los intervalos de confianza para los estudiantes de ingeniería en México*. Tesis doctoral. Universidad de Granada.
- Vallecillos, A. (1994). *Estudio teórico-experimental de errores y concepciones sobre el contraste estadístico de hipótesis en estudiantes universitarios*. Tesis Doctoral. Universidad de Granada.

- Vera, O., Díaz, C. y Batanero, C. (2011). Dificultades en la formulación de hipótesis estadísticas por estudiantes de Psicología. *Unión*, 27, 41-61.
- White, A. L. (1980). Avoiding errors in educational research. En R. J. Shumway (Ed.), *Research in mathematics education* (pp. 47 – 65). Reston, Va: National Council of Teachers of Mathematics.