

UNIVERSIDAD DE GRANADA

Departamento de Didáctica de la Matemática



**COMPRENSIÓN INTUITIVA Y APRENDIZAJE FORMAL DE LAS
TABLAS DE CONTINGENCIA EN ALUMNOS DE PSICOLOGÍA**

Gustavo R. Cañadas de la Fuente

Directores: Carmen Batanero y Antonio Estepa

Editor: Editorial de la Universidad de Granada
Autor: Óscar A. Rodríguez Domínguez
D.L.: GR F/1-2013
ISBN: 978-84-9028-1-1

COMPRENSIÓN INTUITIVA Y APRENDIZAJE FORMAL DE LAS
TABLAS DE CONTINGENCIA EN ALUMNOS DE PSICOLOGÍA

TESIS DOCTORAL

MEMORIA realizada bajo la dirección de la
Dra. Carmen Batanero Bernabeu, y
Dr. Antonio Estepa Castro
que presenta D. Gustavo R. Cañadas de la Fuente
para optar al grado de Doctor

Fdo: Gustavo R. Cañadas de la Fuente

Vo. Bo.

Dra. Carmen Batanero Bernabeu

Dr. Antonio Estepa Castro

Trabajo realizado en el marco de la beca FPU-AP2009-2807, Proyecto EDU2010.14947 (MICINN y FEDER) y Grupo PAI FQM126 (Junta de Andalucía).

AGRADECIMIENTOS

Este trabajo no habría sido posible sin una serie de personas a los que quiero agradecer enormemente la dedicación prestada, más concretamente, y en primer lugar, a los doctores que me han dirigido, supervisado y que se han preocupado de hasta el último detalle para que este proyecto llegase a buen puerto. Ellos han sido un faro hacia el que miraba y me dirigía en este camino, a veces tan complicado, que supone llegar a ser doctor.

Así mismo, debo y quiero agradecer a las siguientes personas, su implicación en las diferentes fases de elaboración de esta tesis:

A cada una de las personas que han leído este compendio de palabras, pues sin ellas no habría sido posible darlo por terminado.

A los miembros del grupo de investigación, que me prestaron su tiempo incondicionalmente cuando necesité su ayuda.

A mis compañeros del Dpto de Didáctica de la Matemática, por su apoyo y ayuda con algunas sugerencias que resultaron de interés para este trabajo.

A mis amigos y gente más cercana, que saben cuántas veces tuve que ausentarme de su grata compañía para llegar hasta aquí y escribir este trabajo.

A todos aquellos que se han preocupado desinteresadamente por mí en algún momento de esta etapa de mi vida, que han compartido los buenos y malos ratos, que hicieron un esfuerzo para facilitar el mío o por una simple sonrisa que me han dedicado siempre que la he necesitado.

A Pepe Camacho, por dejarme utilizar una de sus creaciones, para enmarcar este trabajo.

A mi padre y mi hermano, por su interés y preocupación por mí, en todo este tiempo.

Y a mi madre, por estar ahí, siempre, en ésta y en todas las aventuras de mi vida. Gracias.

COMPRENSIÓN INTUITIVA Y FORMAL DE TABLAS DE CONTINGENCIA POR ESTUDIANTES DE PSICOLOGÍA

1. El problema de investigación	11
1.1. Introducción	11
1.2. Importancia del tema en la formación del psicólogo	12
1.2.1. Lectura y construcción de tablas, como componente de la cultura estadística	12
1.2.2. Análisis de tablas de contingencia en el trabajo e investigación psicológica	15
1.2.3. Tablas de contingencia en los estudios de psicología	17
1.3. Marco teórico	19
1.3.1. Introducción	19
1.3.2. La actividad matemática y objetos ligados a ellas	19
1.3.3. Facetas o dimensiones de los objetos matemáticos	22
1.3.4. Relaciones entre objetos: función semiótica	24
1.3.5. Competencia, comprensión y aprendizaje	25
1.3.6. Evaluación en el enfoque ontosemiótico	26
1.3.7. Criterios de idoneidad didáctica	27
1.4. Tablas de contingencia como objeto matemático	28
1.4.1. Problemas que dan origen a las tablas de contingencia	29
1.4.2. Lenguaje	31
1.4.3. Procedimientos	33
1.4.4. Conceptos y propiedades relacionados con el análisis de tablas de contingencia	38
1.4.5. Argumentos	41
1.5. Software para el análisis de tablas de contingencia	41
1.6. Conclusiones sobre el significado de referencia de la tabla de contingencia	43
1.7. Objetivos del trabajo	45
1.8. Hipótesis iniciales	47
1.9. Estructura del trabajo y características metodológicas generales	51

2. Juicios de asociación y cálculo de probabilidades en tablas de contingencia	55
2.1. Introducción	55
2.2. Asociación, condicionamiento y causación	56
2.2.1. Introducción	56
2.2.2. Elementos de la relación causal	57
2.2.3. Teorías filosóficas sobre causalidad	58
2.2.4. Causalidad desde el punto de vista estadístico	65
2.3. Juicios de asociación	70
2.3.1. Desarrollo evolutivo del concepto de asociación	70
2.3.2. Estrategias en los juicios de asociación	75
2.3.2.1. Investigaciones en psicología	75
2.3.2.2. Sesgos en la detección de la asociación y efecto de las teorías previas	77
2.3.2.3. Investigación en didáctica de la matemática	79
2.3.3. Concepciones sobre la asociación	83
2.3.4. Evolución después de la enseñanza	83
2.3.4.1. Evolución de estrategias en juicios de asociación	84
2.3.4.2. Juicios de asociación usando ordenador	85
2.3.4.3. Evolución de las concepciones	88
2.3.4.4. Actos de comprensión de la asociación	88
2.4. Cálculo de probabilidades a partir de tablas de contingencia	91
2.4.1. Investigación con estudiantes de psicología	92
2.4.2. Trabajos con futuros profesores	93
2.4.3. Otros trabajos	97
2.5. Comprensión del contraste de hipótesis	98
2.6. Conclusiones del estado de la cuestión	106
3. Evaluación inicial de estrategias intuitivas y estimación de la asociación	109
3.1. Introducción	109
3.2. Objetivos	110
3.3. Cuestionario	111

3.3.1.	Selección y adaptación de ítems	111
3.3.2.	Análisis del contenido de los ítems del cuestionario	114
3.3.3.	Prueba piloto y revisión	119
3.4.	Revisión del instrumento	121
3.5.	Descripción de la muestra	122
3.6.	Juicios de asociación	124
3.7.	Estimación del coeficiente de asociación	126
3.8.	Consistencia de los datos	129
3.9.	Teorías previas sobre la asociación	133
3.10.	Análisis de estrategias	134
3.10.1.	Niveles de elaboración de estrategias	136
3.10.2.	Metodología del análisis semiótico de estrategias	140
3.10.3.	Configuraciones cognitivas en las estrategias	141
3.10.4.	Conflictos semióticos	174
3.11.	Conclusiones sobre el estudio inicial de evaluación	180
4.	Diseño de un proceso de estudio de las tablas contingencia para estudiantes de Psicología	183
4.1.	Introducción	183
4.2.	Descripción general del proceso de estudio	184
4.2.1.	Contexto curricular	184
4.2.2.	Estructuración	186
4.2.3.	Teoría sobre la instrucción implícita en el diseño	187
4.2.4.	Distribución temporal	189
4.3.	Análisis de la primera lección. Tablas de contingencia	191
4.3.1.	Objetivos y contenidos	191
4.3.2.	Desarrollo	191
4.3.3.	Significado institucional pretendido	197
4.4.	Análisis de la segunda lección. Asociación estadística, dependencia funcional e independencia	198
4.4.1.	Objetivos y contenidos	198
4.4.2.	Desarrollo	198
4.4.3.	Significado institucional pretendido	204

4.5. Análisis de la tercera lección. Estadístico Chi-cuadrado y contrastes asociados	205
4.5.1. Objetivos y contenidos	205
4.5.2. Desarrollo	206
4.5.3. Significado institucional pretendido	213
4.6. Análisis de la cuarta lección. Medidas de asociación	214
4.6.1. Objetivos y contenidos	214
4.6.2. Desarrollo	215
4.6.3. Significado institucional pretendido	221
4.7. Sesión de prácticas	222
4.7.1. Programa de cálculo Excel	222
4.7.2. Ejercicios de la primera sesión de prácticas	228
4.7.3. Significado institucional pretendido	235
4.8. Conclusiones sobre el significado institucional pretendido	237
5. Evaluación del aprendizaje mediante cuestionario	243
5.1. Introducción	243
5.2. Objetivos y características de los cuestionarios	244
5.2.1. Objetivos	244
5.2.2. Proceso de construcción	245
5.3. Análisis a priori del significado evaluado en el cuestionario	246
5.3.1. Análisis del cuestionario 1	246
5.3.2. Análisis del cuestionario 2	253
5.3.3. Análisis de la prueba final	260
5.4. Síntesis de las evaluaciones mediante cuestionario	264
5.5. Resultados en el cuestionario	267
5.5.1. Resultados en el primer cuestionario	268
5.5.2. Resultados del segundo cuestionario	276
5.5.3. Síntesis de resultados en los cuestionarios	285
5.6. Resultados en la prueba final	298
5.7. Conclusiones	304
6. Evaluación del aprendizaje mediante problemas abiertos	309

6.1.Introducción	309
6.2.Objetivos de la evaluación	310
6.3.Análisis a priori del significado evaluado en los ítems abiertos	311
6.3.1. Análisis del problema A: Contraste de homogeneidad	312
6.3.2. Análisis del problema B: Medidas de asociación en tablas 2x2	314
6.3.3. Análisis del problema C: Medidas de asociación en tablas rxc	316
6.3.4. Análisis del problema D: Cálculo de probabilidades y juicios de asociación	317
6.3.5. Análisis del problema E: Juicios de asociación y justificación	319
6.4.Síntesis de las evaluaciones mediante problemas abiertos	321
6.5.Metodología de análisis de los ítems abiertos	322
6.6.Resultado en el problema A: Contraste de homogeneidad	324
6.6.1. Formulación de hipótesis	324
6.6.2. Estadístico de contraste y probabilidad asociada	330
6.6.3. Toma de decisión	334
6.6.4. Interpretación de resultados	335
6.7.Resultado en el problema B: Medidas de asociación en tablas 2x2	339
6.7.1. Coeficiente Phi de Pearson	339
6.7.2. Riesgo Relativo	342
6.7.3. Razón de Productos Cruzados	344
6.8.Resultado en el problema C: Medidas de asociación en tablas rxc	346
6.8.1. Coeficiente de Contingencia de Pearson	347
6.8.2. V de Cramer	350
6.8.3. Lambda de Goodman y Kruskal	352
6.9.Resultado en el problema D: Cálculo de probabilidades y juicios de asociación	355
6.9.1. Cálculo de Probabilidad	356
6.9.2. Juicio de asociación	359
6.9.3. Estrategia en el juicio de asociación	360
6.10. Resultado en el problema E: Juicio de asociación y justificación	365

6.10.1. Juicio de asociación	365
6.10.2. Signo de la asociación	366
6.10.3. Intensidad de la asociación	367
6.10.4. Justificación del juicio de asociación	367
6.11. Síntesis de resultados en los problemas abiertos	372
6.12. Relación entre conocimiento teórico y resolución de problemas	375
6.13. Conclusiones	376
7. Conclusions	385
7.1. Introduction	385
7.2. Conclusions regarding the general aims	386
7.3. Conclusions regarding the hypotheses	393
7.4. Study contributions and limitations	398
7.5. Ideas for future research	399
Referencias	401
Anexos	
A1. Cuestionario de evaluación de conocimientos previos	A1
A2. Ejercicios propuestos en las sesiones prácticas y evaluaciones	A7
A3. Cuestionarios de evaluación del aprendizaje	A11
A4. Resumen de la observación de la enseñanza	A19
A5. Desarrollo teórico del tema	A39
A6. Descripción del programa de cálculo Excel	A75

INTRODUCTION

In this document we present a research which consists of several phases, all of them focused on the training of professionals in the field of psychology in relation to contingency tables. The choice of this topic is justified by its role in the set of statistical methods, its usefulness in everyday life and professional work, and the many difficulties described in previous research from both psychology (e.g., by Inhelder & Piaget, 1955; Smedlund, 1963, Jenkins & Ward, 1955; Shaklee & Mins, 1982, Perez-Echeverria, 1990 or McKenzie, 2007) and mathematical education (Estepa, 1993; Batanero, Estepa, Godino & Green, 1996; Batanero, Estepa & Godino, 1997; Batanero, Godino & Estepa, 1998).

While that research centered on describing the intuitive strategies and biases of students or professionals in association judgments, in the absence of formal training on contingency tables, our work intends to go a little further and addresses both the initial assessment, as well as the study of the effect of teaching in natural groups of students, while taking into account the educational system restrictions. We also want to evaluate the possible difficulties that remain after the teaching process.

To achieve these objectives, the memory is organized into several distinct parts. The first chapter is devoted to introducing the research problem, justifying the reasons why we focus on psychology students, and on this mathematical object, that we consider complex, despite its simple appearance. We also perform a description of the theoretical framework supporting our work, which is the onto semiotic approach developed by Godino and collaborators. We then perform an epistemic analysis of the mathematical object “contingency table”, by describing the problems that gave rise to this object, and the mathematical language, concepts, methods, properties and arguments required for solving these problems. A brief description of its management with usual software, such as SPSS or Statgraphics, is also included. All of this fixes the reference meaning, and the research objectives and hypotheses. The chapter is finished with a brief description of the work structure and of its general methodological characteristics.

In the second chapter we summarize previous research, in which our study is

supported. We begin by briefly analyzing the concepts of association, conditioning and causation, and their mutual relations, which may be conflicting for some people, as suggested by Diaz and de la Fuente (2005b). We then focus on association judgments, describe its evolutionary development, whose stages were investigated by Inhelder and Piaget (1955), and analyze the strategies used in association judgments by adults, as well as the biases described in the field of psychology.

We also make a detailed study of the research carried out by Estepa and his colleagues (Estepa, 1993; Batanero, Estepa, Godino & Green, 1996; Batanero, Estepa & Godino, 1997; Batanero, Godino & Estepa, 1998), who described conceptions of statistical association and studied their evolution in a teaching experiment; they also described acts of understanding of the concept of association. The chapter is finished with a discussion of research related to computing probabilities from contingency tables.

In the third chapter, we describe an initial evaluation study in a large sample of psychology students. Using four items adapted from Estepa (1993), we analyze the accuracy in the association judgment and the estimation of its strength, the students' previous theories and intuitive strategies, as well as the effect of various task variables on these points. The students strategies are classified from different points of view: First we study the number of cells and the type of frequency used and the processing levels defined by Pérez Echeverría (1990). Then, we perform a semiotic analysis of an example in each strategy, by identifying the mathematical objects and processes involved in the mathematical practices linked to the resolution of the problems posed and the students' semiotic conflicts. The classification of semiotic conflicts is an original contribution of the study.

In the fourth chapter, we analyze a process of study of contingency tables for psychology students, which was designed for this research, and takes into account the most common conflicts found in the initial assessment, as well as those described in previous research, the epistemic analysis presented in Chapter 1 and the curricular constraints. We present an overview of the study process, the curricular context, structure, instructional theory implicit in the design, criteria to select the content and organization of the teaching time. Additionally, we analyze in detail each lesson designed in the study process, stating its objectives, describing the expected development, the mathematical objects used and a synthesis of the intended meaning. We planned for theoretical sessions, where students worked in large group, as well as

and practical sessions, in which they worked in small groups with a software developed for the teaching. We also prepared some materials for the students and some assessment instruments (see Annexes 2-7).

The evaluation of learning is carried out with two types of instrument. Firstly, we analyze the responses to a series of questionnaires consisting of true/false and multiple-choice items (Chapter 5), which allow the quick assessment of learning about a large number of mathematical objects. In this chapter we discuss the questionnaires objectives, the construction process and characteristics. In addition we carry out an a priori analysis of the meaning for each item, to show its validity and epistemic suitability. We present and discuss in detail the responses to the questionnaires, analyze the student' learning in each lesson, and present an overall summary of results (difficulty, discrimination and reliability, total and partial scores).

In Chapter 6, we complete the previous evaluation with an analysis of responses to a number of open-ended problems, which allows a more profound access to the students' understanding, their argumentative ability and the use of the software provided. After describing the objectives, the construction process and the a- priori meaning evaluated by such problems we discuss the methodology of the analysis and present results in each problem. This analyses carried out in chapters 5 and 6 suggest that the learning achieved is quite satisfactory and shows the improvement of intuitive strategies and the global knowledge of the students; however some conflicts remain after teaching.

We think the research work carried out is complex and takes into account previous research. It also allows advancing our knowledge of what students learn in real classrooms, in a study process designed for a specific content. We also provided criteria on areas that can be improved in the future teaching of the subject.

The material designed for teaching and the assessment instruments are other useful contribution to teaching to future students of psychology and can be adapted to other educational contexts. In addition, the information on semiotic conflicts, as well as on the understanding of the association coefficients (a subject with no previous research) complement the previous research. Part of these results have been published in mathematics education journals and conference proceedings.

CAPÍTULO 1.

EL PROBLEMA DE INVESTIGACIÓN

- 1.1. Introducción
- 1.2. Importancia del tema en la formación del psicólogo
 - 1.2.1. Lectura y construcción de tablas, como componente de la cultura estadística
 - 1.2.2. Análisis de tablas de contingencia en la investigación psicológica
 - 1.2.3. Tablas de contingencia en los estudios de psicología
- 1.3. Marco teórico
 - 1.3.1. Introducción
 - 1.3.2. La actividad matemática y objetos ligados a ellas
 - 1.3.3. Facetas o dimensiones de los objetos matemáticos
 - 1.3.4. Relaciones entre objetos: función semiótica
 - 1.3.5. Comprensión, competencia y aprendizaje
 - 1.3.6. Evaluación en el enfoque ontosemiótico
 - 1.3.7. Criterios de idoneidad didáctica
- 1.4. Tablas de contingencia como objeto matemático
 - 1.4.1. Problemas que dan origen a las tablas de contingencia
 - 1.4.2. Lenguaje
 - 1.4.3. Procedimientos
 - 1.4.4. Conceptos y propiedades relacionados con el análisis de tablas de contingencia
 - 1.4.5. Argumentos
- 1.5. Software para el análisis de tablas de contingencia
- 1.6. Conclusiones sobre el significado de referencia de la tabla de contingencia
- 1.7. Objetivos del trabajo
- 1.8. Hipótesis iniciales
- 1.9. Estructura del trabajo y características metodológicas generales

1.1. INTRODUCCIÓN

El desarrollo de la estadística en el siglo XX ha influido en el avance de la ciencia y la sociedad, al proporcionar herramientas metodológicas que pueden ser aplicadas en áreas muy diferentes para analizar la variabilidad, determinar relaciones entre variables, diseñar estudios y experimentos y toma de decisiones en situaciones de incertidumbre (Batanero, 2002).

Ridgway, Nicholson y McCusker (2008), analizan la importancia de la cultura estadística, debida, entre otros hechos al esfuerzo de instituciones como la Unión Europea o la Organización de Naciones Unidas de hacer accesible a los ciudadanos sus estadísticas. Estas estadísticas, con frecuencia manejan representaciones gráficas o datos complejos que se ponen a disposición del público en Internet. Consecuentemente, el interés de que los ciudadanos adquieran un conocimiento estadístico básico es una

necesidad prioritaria (Carrión y Espinel, 2006). A raíz de esta necesidad, la enseñanza de la estadística se ha incorporado de forma generalizada, desde hace unas décadas, en todos los niveles educativos. Ello es debido al carácter instrumental de la estadística, y el valor del desarrollo del razonamiento estadístico en la sociedad de la información (Batanero, 2002).

Además, la Estadística está presente en el curriculum de multitud de carreras universitarias, ya que se aplica a muchos campos de conocimiento, uno de ellos es la de Psicología.

Por otra parte, no es solamente la multitud de ámbitos de conocimiento donde se enseña estadística, sino que desde hace tiempo se demanda nuevos contenidos y nueva pedagogía para los cursos universitarios de iniciación a la Estadística (Moore, 1997)

Estas consideraciones nos han llevado a centrar nuestro trabajo en la Didáctica de la Estadística y más específicamente, en la comprensión de las tablas de contingencia por los estudiantes de psicología. En los apartados siguientes comenzamos resaltando importancia del tema en la formación del psicólogo, presentamos el marco teórico y, realizamos un análisis del objeto “tablas de contingencia”, utilizando algunos elementos de dicho marco teórico. Para finalizar, describimos los objetivos del trabajo, las hipótesis iniciales, su estructura y características metodológicas generales.

1.2. IMPORTANCIA DEL TEMA EN LA FORMACIÓN DEL PSICÓLOGO

En esta sección realizamos un análisis de la importancia que la comprensión de las tablas de contingencia tienen para los estudiantes de psicología apoyándonos en tres argumentos: (a) su papel en la cultura estadística de todo ciudadano; (b) su importancia para la actuación profesional y la investigación en psicología; (c) el tema se incluye en las asignaturas de análisis de datos y es fundamento de otros contenidos en dichas asignaturas.

1.2.1. LECTURA Y CONSTRUCCIÓN DE TABLAS, COMO COMPONENTE DE LA CULTURA ESTADÍSTICA

Las tablas y gráficos estadísticos pueden utilizarse para comunicar información y como instrumento de análisis de datos, así como para retener en la memoria una gran cantidad de información en forma eficiente (Cazorla, 2002). Tienen un papel relevante para organizar describir y analizar los datos; de acuerdo a Wild y Pfannkuch (1999) las podemos considerar como un instrumento de *transnumeración*, pues pueden

proporcionar nueva información no disponible en los datos brutos.

Las tablas y gráficos estadísticos son representaciones semióticas externas, que se utilizan en muchas ciencias; constituyen puente entre los datos experimentales y las formalizaciones científicas y ayudan a determinar las relaciones entre las variables que intervienen en los fenómenos, para poder modelizarlos (Espinel, 2007; González, Espinel & Ainley, 2011). Por tanto, el aprendizaje de los conceptos científicos está ligado al de estas representaciones y al de sus procesos de construcción y transformación, ya que en la enseñanza de las ciencias las tablas y gráficos ayudan a visualizar conceptos y relaciones abstractas (Postigo y Pozo, 2000).

Lectura e interpretación de tablas

De acuerdo a Schield (2006) cualquier persona culta debiera poder leer críticamente las tablas estadísticas que encuentra en la prensa, Internet, medios de comunicación, y trabajo profesional. Esto supone no sólo la lectura literal, sino ser capaz de identificar las tendencias y variabilidad de los datos, y detectar los posibles errores conscientes o inconscientes que puedan distorsionar la información representada. Schield (2000) indica que la lectura de una tabla implica la comprensión de los títulos, subtítulos y los encabezados de filas y columnas; así como el tipo de información incluida. Esto no es sencillo, pues muchas de las tablas que aparecen en la prensa o Internet, combinan diversos tipos de información numérica (frecuencias, porcentajes), clasificada en función en dos o más variables. Tampoco es siempre obvio determinar si el total para una cierta categoría viene representada en una columna, fila, o en los márgenes. Aunque matemáticamente no hay diferencia en representar los porcentajes que elegimos para comparar en la fila o la columna, la facilidad de lectura es diferente, pues al comparar dos porcentajes en la misma columna, tienen un total común, lo que no ocurre si se compara una fila y una columna o dos columnas diferentes.

Por otro lado, el autor indica que en las tablas que leemos en la prensa diaria aparecen proporciones, porcentajes y razones que pueden ser difíciles de comprender o comparar¹. Por ejemplo, a veces es difícil diferenciar la proporción de hombres que fuman (frecuencia doble) con la proporción de hombres entre los fumadores (frecuencia condicional). Los porcentajes pueden referirse a una parte del total: “cada uno de los

¹ Schield (2000) analizó las tablas publicadas en Estados Unidos en 1997, indicando que alrededor del 30% contenían porcentajes, el 10% razones y el 10% otros estadísticos (media, mediana, percentiles).

chicos recibe el 25% de la herencia” o a la razón de un cambio: “Las ventas tuvieron un 20% de incremento”. Las razones pueden ser una parte del todo “La razón de muertes en accidentes de vehículos de motor fue de 16 por 100000 habitantes” o referirse a un intervalo de tiempo: “La razón de velocidad fue de 100 Km por hora. La capacidad de realizar estas comparaciones es parte de la cultura estadística según Schield (2000).

Niveles de lectura

Diversos autores han definido niveles de lectura y comprensión de tablas y gráficos estadísticos, siendo la clasificación más conocida la de Curcio (1989), quien definió los siguientes:

- a. “Leer entre los datos”; consiste en la lectura literal del gráfico o tabla sin interpretar la información contenida en el mismo;
- b. "Leer dentro de los datos": en este nivel se requiere una interpretación e integración de los datos;
- c. "Leer más allá de los datos": realizar predicciones e inferencias a partir de los datos sobre informaciones que no se reflejan directamente en el gráfico o tabla.

Friel, Curcio y Bright (2001) amplían la clasificación anterior definiendo un nuevo nivel “leer detrás de los datos” consistente en valorar críticamente el método de recogida de datos, su validez y fiabilidad, así como las posibilidades de extensión de las conclusiones.

Cuando se considera no sólo la interpretación de los datos, sino también su valoración crítica, los niveles superiores se modifican ligeramente y la categoría “leer detrás de los datos”, puede subdividirse, en función de la capacidad crítica, respecto a la información reflejada en la tabla o gráfico (Aoyama, 2007):

- *Nivel Racional/Literal*. Los estudiantes leen correctamente la tabla o gráfico, incluyendo la interpolación, detección de tendencias y predicción, pero no cuestionan la información, ni dan explicaciones alternativas.
- *Nivel Crítico*. Los estudiantes leen las tablas o gráficos, comprenden el contexto y evalúan la fiabilidad de la información, cuestionándola a veces, pero no son capaces de buscar hipótesis que expliquen la discordancia entre un dato y una interpretación del mismo.
- *Nivel Hipotético*: Los estudiantes leen los gráficos y tablas, los interpretan y evalúan la información, formando sus propias hipótesis y modelos:

La interpretación crítica de tablas y gráficos moviliza diversos conocimientos y es un proceso complejo en el que muchos elementos se ven implicados, entre los que destacan tres aspectos: cognitivo, contextual y afectivo (Monteiro y Ainley, 2004). Monteiro y Ainley (2007) se preocupan de la laguna existente entre la interpretación de gráficos estadísticos en contexto escolar y contextos extra-escolares, como la prensa. En el contexto escolar, se insiste en los conocimientos y procesos estadísticos, prestando poca atención al contexto social del que han sido tomados los datos.

Construcción de tablas

También es posible definir varios niveles de dificultad en la construcción de tablas y gráficos. Por ejemplo Arteaga (2008) propuso una clasificación de los gráficos producidos por futuros profesores cuando trabajan en una tarea abierta, en función de su complejidad semiótica y los niveles de lectura que posibilitaban según la clasificación dada por Friel, Curcio y Bright (2001).

Esta clasificación podría aplicarse a las tablas. El autor diferencia los siguientes niveles: (1) representación de datos individuales; (2) representación de un conjunto de datos, sin llegar a resumir su distribución; (3) representación de una distribución de datos y (4) representación de varias distribuciones sobre un mismo gráfico. Los gráficos y tablas más complejos en su clasificación facilitan la resolución de problemas estadísticos, pero, al mismo tiempo, los errores de interpretación de los mismos crecen con esta complejidad. Es interesante observar que una tabla de contingencia siempre sería una representación de nivel (4) puesto que representa conjuntamente dos variables estadísticas y de ella se deducen varias distribuciones (marginal de cada variable, condicionales y doble).

1.2.2. ANÁLISIS DE TABLAS DE CONTINGENCIA EN EL TRABAJO E INVESTIGACIÓN PSICOLÓGICA

El conocimiento de las tablas de contingencia es también importante para la práctica profesional e investigación de los psicólogos, ya que ellos las usan con frecuencia cuando tratan de extraer y resumir información útil de las observaciones de sus investigaciones.

Las tablas de contingencia son fundamentales, en las actividades de diagnóstico y evaluación psicológica, en las que con frecuencia hay diferentes síntomas que pueden

estar asociados o no con diferentes patologías. La realización de cualquier estudio clínico-epidemiológico pretende poner de manifiesto, al final del mismo, si existe o no asociación entre diferentes variables. Díaz y Gallego (2006) indican que los síntomas o pruebas médicas que se asocian a una enfermedad, de estar presentes, aumentan las probabilidades de dicha afección. Indican que los cálculos exactos son difíciles y sugieren organizar los datos en una tabla de contingencia, similar a la Tabla 1.2.2.1, para facilitar el diagnóstico.

Tabla 1.2.2.1. Ejemplo de tabla de contingencia en una prueba médica

Síntoma	Con enfermedad	Sin la enfermedad	Total
Presente	<i>A</i>	<i>B</i>	<i>A+B</i>
Ausente	<i>C</i>	<i>D</i>	<i>C+D</i>
Total	<i>A+C</i>	<i>B+D</i>	

A partir de esta tabla definen los siguientes indicadores clínicos, que son usuales en la práctica médica y psicológica y se calculan directamente de la tabla de contingencia:

- *A: verdaderos positivos* (número de individuos con la enfermedad y el síntoma presente).
- *B: falsos positivos* (número de individuos sin la enfermedad y el síntoma presente).
- *C: falsos negativos* (número de individuos con la enfermedad y el síntoma ausente).
- *D: verdaderos negativos* (número de individuos sin la enfermedad y el síntoma ausente).

$$\text{Sensibilidad} = \frac{A}{A+C} \times 100 \quad \text{Especificidad} = \frac{D}{B+D} \times 100$$

$$\text{VPP} = \frac{A}{A+B} \times 100 \quad \text{VPN} = \frac{D}{C+D} \times 100$$

Respecto a la investigación psicológica, se emplean usualmente las tablas de contingencia, cuando se utilizan datos cualitativos. Es frecuente el uso de estas tablas, tanto para presentar los datos de las muestras en los estudios, como en el análisis de asociación entre variables o de homogeneidad de muestras (ambas mediante el contraste Chi-cuadrado). En el estudio de la asociación, generalmente se calcula algún coeficiente de asociación, siendo el más habitual el de contingencia y también los que se describirán en el Capítulo 4, al describir el material preparado para la enseñanza del tema. Las tablas de contingencia pueden también realizarse con varias dimensiones o

variables y es posible predecir una de ellas a partir del resto mediante los modelos log-lineales.

1.2.3. TABLAS DE CONTINGENCIA EN LOS ESTUDIOS DE PSICOLOGÍA

Otro motivo que hace relevante el tema es su inclusión en los estudios oficiales de psicología. La estadística se estudia como una materia en la mayoría de planes de estudio de las más diversas profesiones y forma parte de los programas de especialización y posgrado en las más diversas disciplinas. Es, de hecho, casi imposible que un profesional, no haya recibido en su formación al menos un curso de estadística.

En el documento que describe la verificación de los estudio del Grado de Psicología de la Universidad de Granada (ANECA, 2010) un objetivo implícito es proporcionar una base estadística, que capacite para realizar un proceso de resolución de un problema de forma completa:

“El objetivo general del título de grado en Psicología es formar futuros profesionales con los conocimientos científicos necesarios para comprender, interpretar, analizar y explicar los procesos mentales y el comportamiento humano...” (p. 9).

En el Grado de Psicología de la Universidad de Granada, se contemplan 240 créditos distribuidos como aparece en la Tabla 1.2.3.1. Las materias de formación básica y los módulos obligatorios contemplan la formación y adquisición de habilidades necesarias para todo psicólogo. En el caso de las materias optativas se refleja una formación más específica profesional o de investigación.

Tabla 1.2.3.1. Resumen de las materias y distribución en créditos ECTS

Tipo de materia	créditos ECTS
Formación básica	60
Obligatorias	96
Optativas	60
Prácticas externas	18
Trabajo fin de Grado	6
TOTAL	240

En las nuevas titulaciones de grado de Psicología (ANECA, 2005) una de las competencias consideradas necesarias en la mayoría de las universidades europeas en estos estudios es el conocimiento de metodología de investigación y técnicas estadísticas. En la propuesta de la comisión se incluye un Bloque de contenidos formativos sobre *Métodos, diseños y técnicas de investigación en Psicología* con una carga lectiva de 18 créditos, o 450 horas-trabajo del alumno). En sus contenidos

formativos mínimos se incluyen los métodos cuantitativos y cualitativos. Diseños de investigación. Técnicas de Análisis de datos. Fundamentos científicos y distintos diseños de investigación.

Estas asignaturas se interpretan de manera diversa en cada universidad, aunque las tablas de contingencia aparecen en forma implícita o explícita en la mayoría. Por ejemplo, en la Universidad de Granada, entre las asignaturas que forman el bloque de formación básica aparecen dos relacionadas con estadística, denominadas: “Descripción y exploración de datos en Psicología” y “Técnicas de análisis en la investigación psicológica”. Hay además otras dos asignaturas obligatorias donde el conocimiento de estadística se requiere en forma bastante completa: “Métodos y diseños de investigación en Psicología” y “Psicometría”. En estas últimas asignaturas se resumen contenidos de métodos cuantitativos y cualitativos, diseños de investigación, técnicas de análisis de datos y psicometría. Cada una de las asignaturas mencionadas consta de 6 créditos.

En la asignatura “Descripción y exploración de datos en Psicología”, se incluyen los contenidos: tipos de datos, descripción y exploración de un grupo de datos, descripción multivariante de un grupo de datos, generalización e inferencia y formulación y contraste de hipótesis: significación estadística frente a relevancia práctica. Todos estos contenidos están relacionados con el estudio de tablas de contingencia, lo que muestra la importancia de incidir en la mejora del aprendizaje de las tablas de contingencia. La asignatura “Técnicas de análisis en la investigación psicológica” es una continuación de la asignatura anterior donde se profundizan los contenidos sobre contraste de hipótesis e inferencia, incluyendo los datos cualitativos muy relacionados con las tablas de contingencia. Estas dos asignaturas se describirán con detalle al analizar el contexto curricular del experimento de enseñanza (Capítulo 4) y sus guías docentes.

En otras universidades, el tema de Tablas de contingencia también aparece. Por ejemplo, en la Universidad Nacional de Educación a Distancia, en la asignatura “Análisis de Datos I en Psicología”, la encontramos en el tema 8, “Organización y representación conjunta de dos variables”, con los siguientes contenidos: Organización conjunta de datos de dos variables; Distribución conjunta de dos variables categóricas u ordinales; Distribuciones marginales; Distribuciones condicionadas; El concepto de relación entre variables; Representación gráfica conjunta de dos variables cualitativas y ordinales; Representación tridimensional. Asimismo la encontramos en el tema 9. Relación entre variables cualitativas y cuasi-cuantitativas, aparecen los siguientes

contenidos: La prueba de Chi-cuadrado; índices de asociación basados en la prueba de Chi-cuadrado; Índices de correlación para variables ordinales; Relación entre una variable dicotómica y una cuantitativa.

Añadido a ello, la estadística, y en particular las tablas de contingencia, son también base para la comprensión de muchas otras materias y su importancia ha ido aumentando en las Ciencias de la Conducta y, más concretamente, en la Psicología. Como muestra de ello basta consultar las publicaciones recientes de Psicología experimental, Psicología del aprendizaje, Psicología educativa, Psicología social, Psicofísica, etc. Hasta en la Psicología clínica se exige ya un dominio bastante profundo de las técnicas utilizadas en tablas de contingencia. Todo ello refuerza el interés de nuestro trabajo.

1.3. MARCO TEÓRICO

1.3.1. INTRODUCCIÓN

Una vez justificada la importancia del tema en el trabajo y la formación del psicólogo, pasamos a exponer los fundamentos teóricos de la investigación. En este trabajo se utilizan algunas nociones teóricas relacionadas con el Enfoque Ontosemiótico (EOS) desarrollado por Godino y su equipo de colaboradores (Godino, 2002b; Godino, Batanero y Font, 2007; Godino, 2012).

El marco teórico se concreta en tres dimensiones -epistemológica, cognitiva e instruccional-, en el análisis de la enseñanza y el aprendizaje de las matemáticas. Las herramientas teóricas con las que cada una de ellas se aborda son la teoría de los significados institucionales y personales de los objetos matemáticos (Godino y Batanero, 1994, 1998), la teoría de las funciones semióticas (Godino, 2002b; Godino, Batanero y Roa, 2005; Font, Godino y D'Amore, 2000) y teoría de las configuraciones didácticas (Godino, Wilhelmi y Bencomo, 2005; Godino, Contreras y Font, 2006; Font, Planas y Godino, 2010).

En lo que sigue se resumen los principales constructos que utilizamos en nuestro trabajo.

1.3.2. LA ACTIVIDAD MATEMÁTICA Y OBJETOS LIGADOS A ELLA

De acuerdo con Godino (2002b), consideramos la matemática desde una triple perspectiva: (1) como actividad de resolución de problemas, internos o externos a las

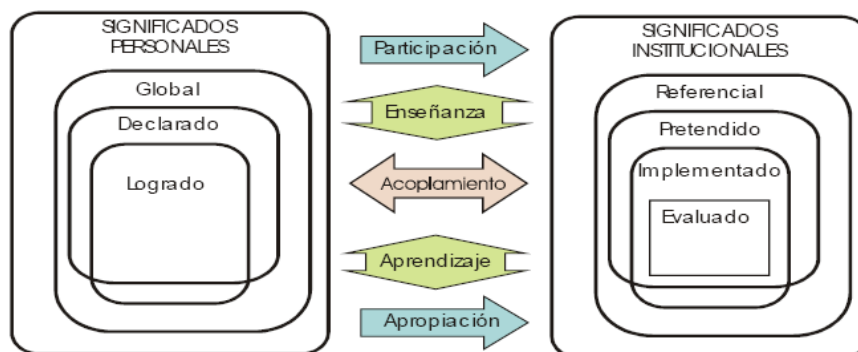
matemáticas; (2) como un lenguaje simbólico propio, en el que se expresan las ideas matemáticas y (3) como un sistema conceptual, con una lógica en su organización. Para operativizar este triple punto de vista de la matemática se toma como noción primitiva la de situación-problemática y a partir de la misma se definen los conceptos de práctica, objeto, tanto personal como institucional, y significado; todos ellos se abordan a continuación.

En este enfoque teórico, se entiende por “práctica matemática” cualquier actuación realizada por alguien y encaminada a la resolución de problemas matemáticos, lo que incluye la comunicación de la solución obtenida, su validación o su generalización a otros contextos y problemas (Godino, 2002b; Godino y Batanero, 1994). En nuestro caso nos interesamos por problemas de los que surge el objeto “tabla de contingencia” y las prácticas involucradas en su resolución.

Las prácticas que se derivan de la actividad matemática pueden ser específicas de una persona o propias de una institución o conjunto de personas interesadas en resolver el mismo tipo de situaciones problemáticas. El hecho de pertenecer a una institución supone el uso de unas prácticas sociales condicionadas por los instrumentos disponibles en la misma así como por sus reglas y modos de funcionamiento (Godino, Batanero y Font, 2007).

En el estudio de las matemáticas interesa considerar los “tipos” de situaciones problemáticas, en relación con los sistemas de prácticas, más que centrarse sólo en una práctica particular de un problema concreto. Así, la respuesta a la pregunta ¿qué es el objeto matemático “tabla de contingencia”?, sería: el sistema de prácticas que realiza determinada persona –significado personal –, o compartidas en una institución –significado institucional-, para resolver un tipo de situaciones-problemas en los que interviene la tabla de contingencia. Es decir, los significados dependen de los contextos sociales y de los sujetos, lo que implica un carácter relativo de tales significados.

Figura 1.3.2.1 Tipos de significados institucionales y personales



Los significados dependen del contexto social y de los sujetos, por tanto, su carácter es relativo. La tipología básica de significados, necesaria en el análisis didáctico de los mismos, aparece resumida en la Figura 1.3.2.1 (Godino, 2002a, p.141). El significado institucional puede ser, “global”, que se refiere a lo que significa el objeto en una institución, “referencial”, relativo al significado que se considera en una enseñanza o investigación, “pretendido”, que indica lo que se pretende enseñar, “implementado”, lo que se logra enseñar, y “evaluado” que hace referencia a la parte que se evalúa.

El significado personal se concreta en significado global (todo lo que un sujeto conoce), evaluado (lo que se puede evaluar inicialmente de su conocimiento), declarado (lo que se consigue evaluar pasada la evaluación) y logrado (parte del conocimiento que está de acuerdo con el significado institucional).

En nuestro trabajo nos interesamos por el significado de la tabla de contingencia dentro de la institución “enseñanza en psicología”. El análisis que realizamos en este capítulo determinará el significado institucional referencial en nuestro trabajo. En el Capítulo 4 se definirá el significado pretendido (en la enseñanza dentro de un curso específico en Facultades de psicología) y en los Capítulos 5 y 6 el evaluado (el que evaluemos con nuestros instrumentos de evaluación).

Configuraciones epistémicas y cognitivas

Cuando se realiza, interpreta y evalúa una práctica matemática, se activa una configuración (o sistema de objetos) constituido por situaciones-problemas, lenguajes, conceptos, proposiciones, procedimientos y argumentos (Font y Godino, 2006; Godino, Batanero y Font, 2007). Estos elementos se definen del modo siguiente:

- *Situaciones-problemas*: aplicaciones extra-matemáticas, ejercicios, problemas, acciones que inducen una actividad matemática. En nuestro caso el problema puede ser el estudio de asociación entre dos variables cualitativas, o el estudio de homogeneidad de varias muestras.
- *Lenguajes*: términos, expresiones, notaciones, gráficos que se utilizan para representar los datos del problema, las operaciones que hacemos con ellos, los objetos matemáticos que se utilizan y la solución encontrada. En el estudio de las tablas de contingencia el lenguaje gráfico tiene un peso muy importante (gráficos de barras, estereogramas,...) aunque también se usa el lenguaje verbal (términos como frecuencia doble o condicional), simbólico (expresiones asociadas) e incluso icónico

al trabajar con ordenador.

- *Conceptos- definición:* En las prácticas que llevan a cabo los estudiantes para resolver un problema matemático (en este caso cuando trabajan con el recurso informático) se usan implícita o explícitamente objetos matemáticos, de los cuáles el alumno ha de recordar o aplicar la definición. Por ejemplo, los estudiantes usarán implícitamente los objetos: tabla de contingencia, probabilidad simple, compuesta y condicional, independencia, asociación, frecuencia simple y compuesta, etc..
- *Proposiciones* o enunciados sobre relaciones o propiedades de los conceptos que igualmente se han de emplear al resolver problemas matemáticos. Por ejemplo, cuando los estudiantes tienen que recordar que si dos variables son independientes el coeficiente de asociación tiene un valor igual a cero
- *Procedimientos:* Serían los algoritmos, operaciones, técnicas de cálculo que los estudiantes han aprendido durante la enseñanza previa y que aplican al resolver el problema. En nuestro caso, los estudiantes usarán varios tipos específicos de contraste de hipótesis, así como otras técnicas más intuitivas.
- *Argumentos:* Serían los enunciados usados para validar o explicar las proposiciones y procedimientos o bien la solución de los problemas. Pueden ser deductivos, inductivos, formales o informales.

Godino, Batanero y Font (2007) proponen la existencia de redes de objetos intervinientes y emergentes de los sistemas de prácticas que relacionan entre sí los seis tipos de categorías anteriores y forman *configuraciones* que se clasifican en:

- *Epistémicas:* cuando se trata de objetos institucionales. La configuración epistémica es el conjunto de objetos matemáticos que intervienen en la resolución de las actividades. Dentro de la configuración epistémica se distingue la previa –objetos que se supone conoce el alumno antes de trabajar en la Unidad Didáctica- y la emergente –lo que se supone que va a aprender-.
- *Cognitivas:* cuando se refieren a objetos personales. Se diferencian de las epistémicas porque los sujetos (por ejemplo estudiantes) podrían construir una configuración de objetos diferente de la institucional.

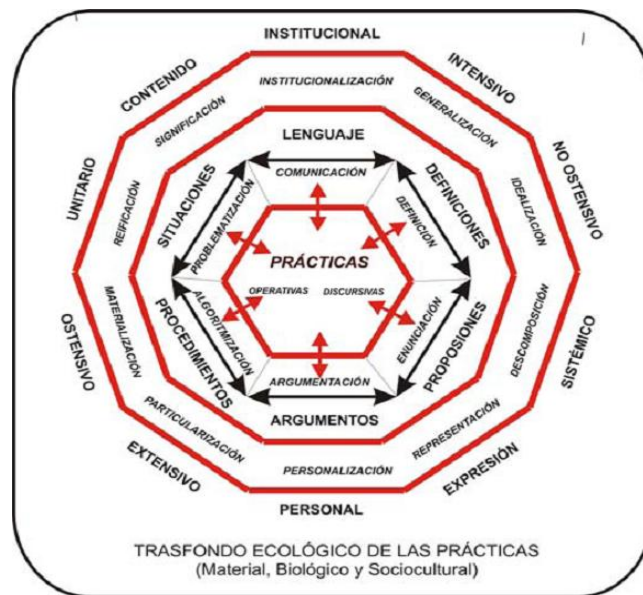
1.3.3. FACETAS O DIMENSIONES DE LOS OBJETOS MATEMÁTICOS

Como vemos en la Figura 1.3.3.1 cada uno de estos objetos tiene asociado un proceso matemático: problematización, comunicación, definición, enunciación,

argumentación y algoritmización (Godino, Batanero y Font, 2007). Además estos autores contemplan los objetos anteriores desde diferentes facetas duales:

- *Personal – institucional*. Si los sistemas de prácticas son compartidas en el seno de una institución, los objetos emergentes se consideran “objetos institucionales”, mientras que si estos sistemas son específicos de una persona se consideran como “objetos personales” (Godino y Batanero, 1994).

Figura 1.3.3.1. Objetos, facetas y procesos que intervienen en las prácticas matemáticas



- *Ostensivo – no ostensivo*. Se entiende por ostensivo cualquier objeto que se puede mostrar a otro, es decir, materializado de alguna forma.
- *Expresión – contenido*. Antecedente y consecuente de cualquier función semiótica (que definimos en la siguiente sección).
- *Extensivo – intensivo*. Un objeto que interviene como un caso particular (un ejemplo específico sería una tabla de contingencia específica o una frecuencia relativa doble) y una clase más general, como “una tabla de contingencia general o una frecuencia relativa doble $\frac{a}{a+b+c+d}$ ”.
- *Unitario – sistémico*. En algunas circunstancias los objetos matemáticos se consideran unitarios (un dato) y en otros como compuestos de varios objetos (la distribución de datos).

También en la Figura 1.3.3.1 vemos que se consideran una serie de procesos asociados a estas dualidades: institucionalización – personalización; generalización – particularización; análisis/descomposición – síntesis/reificación; materialización/concreción – idealización/abstracción; expresión/representación – significación.

En nuestro trabajo utilizaremos la clasificación de tipos de objetos al analizar la enseñanza planificada (Capítulo 4) y los instrumentos de evaluación (Capítulos 5 y 6). También en el análisis semiótico de estrategias incluido en el Capítulo 3.

1.3.4. RELACIONES ENTRE OBJETOS: FUNCIÓN SEMIÓTICA

Otro aspecto de interés en este trabajo es el de función semiótica (Godino, 2002b; Godino, Batanero y Font, 2007). Un objeto matemático, por ejemplo, una tabla de contingencia, al expresarse por escrito o verbalmente, lleva a establecer una función semiótica. Los autores proponen bajo tal denominación, la correspondencia entre un antecedente y un consecuente, establecida por un sujeto, ya sea persona o institución. Sirve para resaltar los procesos de interpretación que se llevan a cabo en la actividad matemática y en los que a veces pueden aparecer desajustes (conflictos) de interpretación entre alumnos y profesor.

La correspondencia entre antecedente y consecuente suele estar implícita, aunque en ocasiones se incluyen instrucciones para la interpretación de los objetos matemáticos. Si las interpretaciones realizadas por los alumnos no son las esperadas por el profesor, se produce un *conflicto semiótico*. El antecedente (expresión) y el consecuente (significado) de una función semiótica no se limitan a conceptos y al lenguaje asociado sino que abarcan toda la ontología de objetos matemáticos u organizaciones de éstos en entidades más complejas, como sistemas conceptuales o teorías. Así, cada texto se puede descomponer en partes, en las que hay implícitos diferentes objetos matemáticos. Mediante el análisis semiótico se intentará poner de manifiesto estos objetos, así como las dificultades y conflictos semióticos que los estudiantes pueden tener al utilizarlos.

Godino (2002b) resalta la importancia del análisis semiótico, que se centra en los objetos y procesos que intervienen en las prácticas matemáticas o emergen de ellos y permite describir la complejidad onto-semiótica de dichas prácticas y ayudar a identificar conflictos semióticos potenciales. Nosotros utilizaremos el análisis semiótico en el Capítulo 3, para clasificar y analizar las estrategias intuitivas de los estudiantes y detectar sus conflictos semióticos al abordar el curso.

1.3.5. COMPETENCIA, COMPRENSIÓN Y APRENDIZAJE

Si queremos analizar la comprensión que tienen un grupo de sujetos de un cierto objeto matemático, es necesario dar respuesta a las preguntas siguientes: ¿qué entendemos por comprensión de un cierto objeto matemático y qué y cómo hacer para lograr dicha comprensión (Godino, 1996). El autor propone un modelo de comprensión en el que considera dos ejes: uno descriptivo, relativo a los aspectos o componentes de los objetos de los que queremos evaluar la comprensión y otro procesual que incluye las fases o niveles necesarios en la consecución de una adecuada comprensión.

Godino (2002a) indica la conveniencia de atribuir un significado distinto y complementario a las nociones de competencia y comprensión matemática, sugiriendo que existen otros usos del término competencia, sobre todo, en contextos de innovación curricular; el sentido en que él lo emplea es referido al complejo cognitivo que comprende aspectos operativos y discursivos del conocimiento matemático.

Godino, Batanero y Font (2007) sugieren que hay dos maneras de entender la "comprensión": como proceso mental o como competencia. Los enfoques cognitivos en la Didáctica de las Matemáticas, en el fondo, entienden la comprensión como "proceso mental". Los posicionamientos pragmatistas del EOS, en cambio, llevan a entender la comprensión básicamente como competencia (se considera que un sujeto comprende un determinado objeto matemático cuando lo usa de manera competente en diferentes prácticas).

Puesto que los autores dan a las funciones semióticas un papel esencial en las prácticas matemáticas, también proporcionan una definición de comprensión también en términos de funciones semióticas. Godino (2003) define la comprensión de un objeto O por parte de un sujeto X (individuo o institución) en términos de las funciones semióticas que X puede establecer, en las que se pone en juego O (como expresión o contenido). Cada función semiótica implica un acto de interpretación y constituye un conocimiento. De ello se deduce una variedad de tipos de conocimientos en correspondencia con la diversidad de funciones semióticas que se pueden establecer entre las diversas entidades introducidas en el modelo.

Por otro lado, en la Figura 1.4.3.1, aparecen representadas las relaciones entre enseñanza y aprendizaje en el EOS. En dicho marco, el aprendizaje se entiende como el acoplamiento progresivo entre los significados personales e institucionales. La enseñanza supone la participación del estudiante en la comunidad de prácticas fijada por la institución, en donde la finalidad es que se produzca el aprendizaje, que implica la

adquisición por parte del estudiante de dichos significados institucionales (Godino, 2003).

En nuestro trabajo analizaremos, por un lado la comprensión intuitiva sobre las tablas de contingencia en el estudio de evaluación realizado en el Capítulo 3, con estudiantes que no han estudiado formalmente el tema. Asimismo estudiaremos el aprendizaje formal adquirido, entendiendo por ello el conjunto de prácticas institucionales adquiridas por el estudiante, como consecuencia de un proceso de estudio matemático del tema. En dicho estudio se introducirán prácticas tales como el uso del contraste Chi-cuadrado y el cálculo de medidas de asociación que difícilmente podría desarrollar el estudiante por sí solo. De ahí la diferencia que establecemos entre comprensión intuitiva y aprendizaje formal.

1.3.6. EVALUACIÓN EN EL ENFOQUE ONTOSEMIÓTICO

Godino (1996) resalta la dificultad de la observación directa de la comprensión personal de un determinado objeto, y propone en su lugar la observación directa de las prácticas personales (significado personal). La evaluación de la comprensión sería el estudio de la correspondencia entre significados personales e institucionales.

Godino, Contreras y Font (2006) indican que, ya que el significado de los objetos matemáticos se concibe como el "sistemas de prácticas operativas y discursivas ligadas a campos de problemas", la evaluación de los significados personales que construyen los alumnos en el proceso de estudio, ha de partir de dichas prácticas.

La evaluación, por tanto, depende de la institución desde donde se lleva a cabo, debido a que es ésta la que determina si un sujeto "comprende" el significado de un objeto. Las prácticas observables son los indicadores empíricos que permiten realizar dicha evaluación. En esta investigación, en los capítulos 5 y 6, seguiremos este sistema para evaluar los conocimientos matemáticos, sobre las tablas de contingencia, que tienen los estudiantes de primero de psicología. Para ello, observaremos sus prácticas personales cuando realizan distintas tareas de evaluación.

Dada la complejidad de los significados personales sobre un objeto no pueden ser abarcados en un solo instrumento de evaluación. En este sentido, Batanero y Díaz (2005), recuerdan los procesos de inferencia que intervienen en un proceso de evaluación y que condicionan la generalizabilidad de los resultados. La evaluación supondría un muestreo de estudiantes, preguntas y circunstancias, que afectan en cada uno de los casos a la posibilidad de inferencia.

Una evaluación con garantías debe poner en juego una muestra representativa de las prácticas que constituyen el significado sistémico de los mismos en el seno de un contexto institucional dado (Godino, 1996). Más aún, si las tareas son suficientemente representativas (para evaluar las unidades de contenido definidas), podríamos hacer una inferencia sobre lo que cada alumno de la muestra sería capaz de hacer y decir en otras tareas relacionadas con el concepto. Finalmente, los alumnos particulares son una muestra (que suponemos representativa) de otros estudiantes, por lo que podremos generalizar a otras poblaciones. En este sentido, Batanero y Díaz (2005) amplían el concepto *idoneidad* y sus tipos, que comentamos en el siguiente apartado, al caso de la evaluación. En nuestro trabajo, tendremos en cuenta dichos criterios.

1.3.7. CRITERIOS DE IDONEIDAD DIDÁCTICA

Los aspectos teóricos previos se han de complementar con el concepto de “Idoneidad Didáctica” del proceso instruccional y para ello, se deben combinar, en mayor o menor grado las seis componentes siguientes (Godino, Contreras y Font, 2006; Godino, Bencomo, Font y Wilhelmi, 2006; Godino, Wilhelmi y Bencomo, 2005; Godino, Font y Wilhelmi; 2008, Godino, 2011).

- *Idoneidad epistémica*: es el grado de representatividad de los significados institucionales implementados (o pretendidos), respecto de un significado de referencia. Sería si los significados de los objetos presentes en un recurso instruccional son adecuados desde el punto de vista matemático.
- *Idoneidad cognitiva*: grado en que los significados pretendidos/ implementados son asequibles a los alumnos, así como si los significados personales logrados por los alumnos son los significados pretendidos por el profesor.
- *Idoneidad interaccional*: Grado en que la organización de la enseñanza permite identificar conflictos semióticos y resolverlos durante el proceso de instrucción.
- *Idoneidad mediacional*: Disponibilidad y adecuación de los recursos necesarios para el desarrollo del proceso de enseñanza-aprendizaje.
- *Idoneidad emocional*: Interés y motivación del alumnado en el proceso de estudio.
- *Idoneidad ecológica*: grado en que el proceso que se estudia se ajusta al entorno; es decir, al proyecto educativo del centro o la sociedad en que se encuentra.

Esta idea será utilizada en el capítulo de conclusiones, para analizar la enseñanza

realizada y los resultados de la evaluación de la misma. Asimismo se aplicará para estudiar la idoneidad del cuestionario de evaluación construido.

En la siguiente sección aplicamos algunas de las ideas anteriores al análisis del objeto matemático “tabla de contingencia”. Específicamente tratamos de identificar los diversos tipos de objetos involucrados en el trabajo con las mismas. La finalidad es lograr una primera aproximación al significado institucional referencial de la tabla de contingencia en nuestro trabajo.

1.4. TABLAS DE CONTINGENCIA COMO OBJETO MATEMÁTICO

En muchos estudios estadísticos se suele estar interesado en más de un carácter latente en la población. Esto implica querer dar respuesta a la pregunta de si existe relación entre ambas variables. En ciertos casos se puede dar una fórmula que exprese una en función de la otra; son los fenómenos deterministas. En los casos en que no es posible dar esta fórmula, pero las dos variables son numéricas, se puede representar pares de valores de las dos variables, mediante puntos en un sistema cartesiano. Entonces esta representación la llamamos nube de puntos o diagrama de dispersión y nos sirve para observar relaciones directas o inversas, y creando aproximaciones por rectas o parábolas por ejemplo.

Otras veces, alguna de las variables no es numérica. Por ello, en el análisis de la relación de dos variables pueden existir tres tipos de estudios, según el tipo de variables implicadas (Batanero y Díaz, 2008):

- *Dos variables o más cualitativas*, donde abordaremos el problema mediante tablas de contingencia. El estudio de la asociación entre dos variables cualitativas puede generalizarse a varias variables, que pueden ser estudiadas, por ejemplo, por los modelos logarítmico lineales.
- *Dos variables cuantitativas*, donde si están relacionadas hablaremos de correlación y regresión, esta última puede ser lineal o curvilínea. En el caso de varias variables cuantitativas se puede estudiar la correlación múltiple o parcial y la regresión múltiple o multivariante.
- *Una variable cualitativa y la otra cuantitativa*, que se puede tratar de varias maneras diferentes, como usando el test t de diferencia de medias o analizando la diferencia de las dos medias. La generalización a varias variables conduce a los modelos de Anova, Ancova y Manova.

En este trabajo nos centramos específicamente en el estudio de la relación entre

dos variables cualitativas, es decir, las tablas de contingencia. A continuación analizamos este objeto matemático.

1.4.1. PROBLEMAS QUE DAN ORIGEN A LAS TABLAS DE CONTINGENCIA

En gran cantidad de estudios estadísticos recogemos datos de cada individuo sobre dos variables estadísticas. Usualmente llamamos: X , a una de las variables con valores x_1, x_2, \dots, x_r ; y la otra variable llamamos Y , con valores y_1, y_2, \dots, y_c . A partir de estos datos bivariantes se presentan varios tipos de problemas, que originaran, primero el concepto de Tabla de Contingencia y otros relacionados con ellas, así como diversos procedimientos y representaciones.

P1. *Resumir la información obtenida de un conjunto de observaciones bivariantes de modo que se visualicen posibles relaciones entre las variables.* El primer problema surge al tratar de representar en forma condensada, un listado de datos bivariantes. La forma habitual de ordenar la información puede ser en forma de *tabla de doble entrada* (también llamada tabla de contingencia). Para construirla, situamos los valores de las variables en la primera fila y columna de la tabla y en cada casilla de frecuencia, cuando hay pares repetidos (Figura 1.4.1.1).

De este modo comenzamos extendiendo el concepto de frecuencia absoluta y relativa del caso bivalente. Siendo f_i la *frecuencia marginal* de la variable x_i ; f_j la *frecuencia marginal* de la variable y_j ; f_{ij} la *frecuencia absoluta* con la que aparece el par (x_i, y_j) ; h_{ij} la *frecuencia relativa* de este mismo par de valores; y siendo n el tamaño de la muestra.

Figura 1.4.1.1. Tabla de doble entrada

	y_1	...	y_j	...	y_c	
x_1						$f_{1.}$
x_2						$f_{2.}$
...						...
x_i						$f_{i.}$
...						...
x_r						$f_{r.}$
	$f_{.1}$...	$f_{.j}$...	$f_{.c}$	n

P2. *Cálculo de probabilidades asociadas a valores específicos de una o las dos variables.* Una vez construida una tabla de contingencia, podemos obtener las probabilidades puntuales de cada una de las variables o también es posible tener la

probabilidad de ocurrencia simultanea de un caso de X y otro de Y (Figura 1.4.1.2). Además es inmediato el cálculo de las probabilidades condicionadas con sólo dividir la probabilidad conjunta entre la probabilidad a la que condicionamos. Es decir:

$$P(X=x_i/Y=y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}$$

Una vez organizada la tabla, cuando tratamos de estudiar la relación entre dos variables cualitativas, aparecen dos nuevos problemas, dependiendo del tipo de relación que se quiera analizar. Llevarían a dos tipos de contraste de hipótesis principalmente: *contraste de homogeneidad y contraste de independencia*.

Figura 1.4.1.2. Probabilidades en la tabla de doble entrada

	Y_1	...	Y_j	...	Y_c	
X_1	$P(X=x_i, Y=y_j)$					$P(X=x_1)$
X_2						$P(X=x_2)$
...						...
X_i						$P(X=x_i)$
...						...
X_r						$P(X=x_r)$
	$P(Y=y_1)$...	$P(Y=y_j)$...	$P(Y=y_c)$	1

P3. *Contraste de homogeneidad*. A veces la variable X representa una población subdividida en r subpoblaciones, donde hemos tomado una muestra de cada subpoblación, y cada individuo lo clasificamos con una variable Y . Siendo p_{ij} una proporción que, en la población x_i tiene $Y=y_j$ de valor, queremos realizar una inferencia para ver si, a partir de las muestras observadas se puede inferir que las subpoblaciones son semejantes. Para llevar a cabo esta inferencia, se realiza un contraste de homogeneidad, siendo las hipótesis, nula y alternativa las siguientes (Batanero y Díaz, 2008):

$$H_0 \equiv p_{1j} = p_{2j} = \dots = p_{mj} \text{ para todo } j.$$

$$H_1 \equiv \text{al menos dos de estas proporciones son diferentes.}$$

Con el resultado del contraste sabremos si las muestras provienen de poblaciones con iguales o diferentes distribuciones de probabilidad; generalmente lo hacemos para combinar resultados si los datos son homogéneos.

P4. *Contraste de independencia*. En este caso nos interesa la asociación o

relación existente entre dos variables de una misma población. Diremos que las variables cualitativas son *independientes* cuando la probabilidad de una categoría de una variable no está influenciada por la categoría de la otra variable en un mismo individuo. Podemos tener tres tipos de tablas: independencia total, asociación perfecta y asociación parcial.

P5. *Definir una medida de la intensidad de asociación entre las variables.* En el caso de que en el problema P4 hayamos llegado a la conclusión de que las variables están asociadas, se plantea un nuevo problema, consistente en definir una medida del grado o intensidad de la asociación entre las mismas.

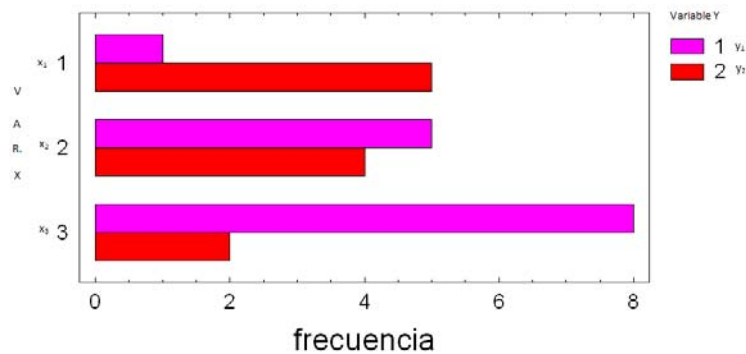
P6. *Ajuste de un modelo matemático que permita estimar la frecuencia de aparición de cada categoría de una variable, dada una categoría de la otra.* En el caso de que en el problema P5, la asociación encontrada sea suficientemente fuerte, se plantea el problema de encontrar un modelo que permita predecir la variable dependiente en función de la independiente, es decir una generalización del problema de regresión numérica. Mediante los modelos logarítmicos lineales se resolverá este problema 6.

1.4.2. LENGUAJE

Existen ligados al estudio de las tablas de contingencia gran variedad de representaciones, verbales, simbólicas, tabulares y gráficas. Las representaciones verbales incluyen todas las palabras utilizadas para referirse a la tabla y sus componentes, como “tabla de contingencia”, “tabla cruzada”, “tabla de doble entrada”. Las representaciones simbólicas y tabulares se han presentado en los apartados anteriores. También hay una gran variedad de gráficos con los que mostrar datos presentados en tablas de contingencia. A continuación mostraremos algunos.

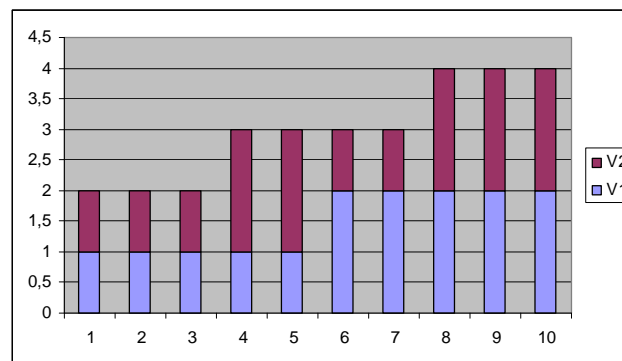
R1. *Diagrama de barras adosadas.* Este sistema de representación nos muestra con barras colocadas horizontalmente o verticalmente la frecuencia de cada casilla del interior de la tabla de frecuencias (Figura 1.4.2.1). Para cada valor de la variable X se presentan conjuntamente la frecuencia de cada valor de Y condicionado a este valor de X . Al estar situadas unas debajo de otras permite la comparación rápida entre las variables y dentro de cada variable. También se podría construir el diagrama de barras en porcentajes, bien absolutos (respecto al total de la tabla), bien relativos (respecto a cada valor de la variable X).

Figura 1.4.2.1. Diagrama de barras adosadas



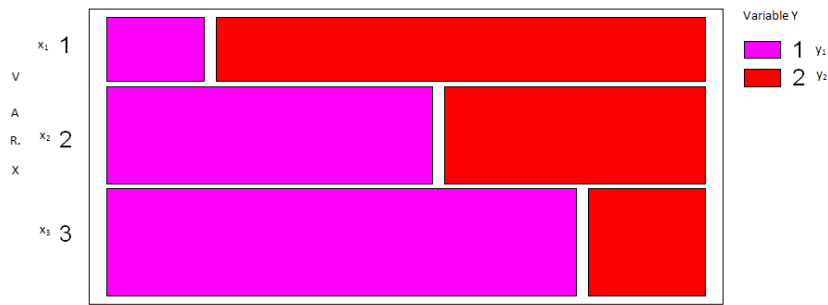
R2. *Diagrama de barras apilado.* Este gráfico muestra las barras unas sobre otras para representar la frecuencia con la que aparece. Compara, entre categorías, el aporte de cada valor al total (Figura 1.4.2.2). También se podría elaborar el diagrama de barras en porcentajes, bien absolutos (respecto al total de la tabla), bien relativos (respecto a cada valor de la variable X). En esta representación se visualizan mejor las frecuencias condicionales que en el diagrama de barras adosado.

Figura 1.4.2.2. Diagrama de barras apilado



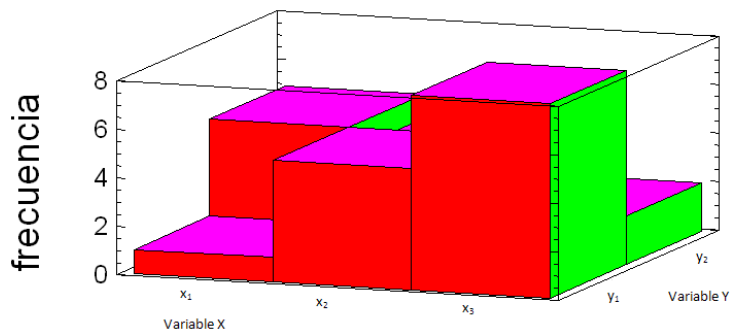
R3. *Gráfico de Mosaico.* Representa mediante áreas la frecuencia de cada celda de la tabla de contingencia (Figura 1.4.2.3). Se divide primero el eje Y en segmentos proporcionales a la frecuencia relativa de cada categoría en filas, es decir, en el eje Y tenemos representadas las frecuencias relativas marginales de la variable en filas. En cada rectángulo resultante se divide el eje X en partes proporcionales a la frecuencia condicional de las columnas de la tabla respecto a la fila dada. El gráfico visualiza las frecuencias marginales de las filas y las frecuencias condicionales de columnas respecto a cada una de las filas de la tabla (Batanero y Díaz, 2008).

Figura 1.4.2.3. Gráfico de Mosaico



R4. *Gráfico tridimensional.* Este gráfico tridimensional coloca en el ancho y largo las variables X e Y, y la altura representa la frecuencia de cada celda (Figura 1.4.2.4). Se puede construir también utilizando frecuencias relativas o porcentajes.

Figura 1.4.2.4. Gráfico tridimensional



1.4.3. PROCEDIMIENTOS

No hay ningún método que cubra todos los problemas planteados en el apartado 1.4.1 y para todos los tipos de tablas de contingencia. Así que estudiaremos algunos de los métodos más usuales, que podemos clasificar en varios apartados.

PR1.1. *Contraste de homogeneidad.* Los más usuales son los basados en la distribución Chi-cuadrado. Para realizarlo utilizamos el estadístico χ^2 como medida de las desviaciones entre la muestra y la población, que viene dado por la siguiente expresión (Ruiz-Maya y Martín, 2002):

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$$

Siendo e_{ij} las frecuencias estimadas considerando la hipótesis nula cierta, es decir:

$$e_{ij} = n_j p_i = n_j \frac{n_{i.}}{n}$$

El estadístico de contraste queda:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(f_{ij} - \frac{n_{i.}n_{.j}}{n})^2}{\frac{n_{i.}n_{.j}}{n}} \text{ que se distribuye con una distribución } \chi^2_{(r-1)(c-1)}$$

Podemos tener dos situaciones (Batanero y Díaz, 2008):

- Si el valor experimental $\chi^2 < \chi^2_{1-\alpha}$, aceptamos la hipótesis nula de que las muestras provienen de poblaciones con iguales distribuciones de probabilidad, para un nivel de significación α (para el Problema P4).
- Si el valor experimental $\chi^2 \geq \chi^2_{1-\alpha}$, no podemos aceptar la hipótesis nula de que las muestras provienen de poblaciones con iguales distribuciones de probabilidad, para un nivel de significación.

PR1.2. Otra manera de calcular la diferencia de subpoblaciones en un caso de homogeneidad es usando “frecuencias esperadas en el caso de ser cierta H_0 ”, para lo cual utilizamos la siguiente expresión:

$$e_{ij} = \frac{f_i f_j}{n}$$

Mediante el estadístico Chi-cuadrado, presente a continuación, tenemos una medida de diferencia entre las frecuencias esperadas y las observadas.

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \rightarrow \chi^2_{(n-1)(m-1)}, \text{ siendo } n \text{ y } m \text{ los tamaños. Podemos tener dos}$$

situaciones (Batanero y Díaz, 2008):

- Si el valor experimental $\chi^2 < \chi^2_{1-\alpha}$, aceptamos la hipótesis nula, para un nivel de significación α .
- Si el valor experimental $\chi^2 \geq \chi^2_{1-\alpha}$, no podemos aceptar la hipótesis nula, para un nivel de significación α .

PR1.3. *Contraste para tablas 2x2*. En el caso particular de tablas 2x2, se puede utilizar para el cálculo del estadístico la siguiente forma:

$$\chi^2 = \frac{(f_{11}f_{22} - f_{12}f_{21})^2 n}{f_{1.}f_{.1}f_{2.}f_{.2}}$$

PR1.4. *Prueba exacta de Fisher.* En todos los casos anteriores, el valor del estadístico se puede aproximar mediante la distribución normal para muestras grandes. Pero, si todas o una parte importante de las frecuencias son pequeñas, en el caso de tablas 2x2 se puede calcular la probabilidad exacta de obtener las frecuencias observadas usando la distribución hipergeométrica. En general se aconseja no usar el contraste Chi-cuadrado a menos que las frecuencias en todas las celdas sean superiores a 5, aunque algunos estadísticos sugieren que basta que todas las frecuencias sean mayores que 1 si el 80% de las celdas tienen frecuencias superiores a 5 (Lipschutz y Schiller, 1999).

PR1.5. *Contraste de independencia.* Para realizar el contraste de independencia (Lipschutz y Schiller, 1999), tomamos como hipótesis nula que las variables son independientes, y realizamos una estimación de p_i y p_j ya que son desconocidos, de la siguiente forma:

$$p_i = \frac{f_{i.}}{n} = \frac{f_{i1} + f_{i2} + \dots + f_{ic}}{n} \quad \text{y} \quad p_j = \frac{f_{.j}}{n} = \frac{f_{1j} + f_{2j} + \dots + f_{rj}}{n}$$

El estadístico de contraste queda

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(f_{ij} - np_i p_j)^2}{np_i p_j} \rightarrow \chi^2_{(r-1)(c-1)}$$

Podemos tener dos situaciones (Batanero y Díaz, 2008):

- Si el valor experimental $\chi^2 < \chi^2_{1-\alpha}$, aceptamos la hipótesis nula de que las variables son independientes, para un nivel de significación α .
- Si el valor experimental $\chi^2 \geq \chi^2_{1-\alpha}$, no podemos aceptar la hipótesis nula de que las variables son independientes, para un nivel de significación α .

PR2. *Medidas de asociación.* En el caso de rechazar la independencia, tendríamos que calcular cuánto están relacionadas o asociadas las variables. Se trata de coeficientes que indiquen el grado de asociación entre las dos variables y en algunos casos su signo. A continuación analizamos los más usuales.

PR2.1. *Medidas de asociación en datos nominales (tabla 2x2)*. Para medir la asociación en este caso (Tabla 1.4.3.1) Pearson nos dio una medida cuya magnitud no depende del tamaño n de la muestra y cuyos valores están en el intervalo (0,1). Cuanto más próximo a 0 mayor es la independencia y cuanto más cerca de 1 mayor es la asociación (Ato y López, 1996). Viene dado por la siguiente expresión:

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

Tabla 1.4.3.1. Caso para medir asociación por Pearson

	B	\bar{B}	Total
A	f_{11}	f_{12}	$f_{1.}$
\bar{A}	f_{21}	f_{22}	$f_{.2}$
Total	$f_{.1}$	$f_{.2}$	n

También podemos calcular en este tipo de tablas (tablas 2x2) el riesgo relativo, correspondiente a cuánto más probable es la presencia de A entre los individuos con la característica B o sin ella, o la *razón de productos cruzados* que es una razón de cocientes y vienen dados por las siguientes expresiones (Ato y López, 1996):

- Riesgo relativo = $RR = \frac{P(A | B)}{P(A | \bar{B})} = \frac{f_{11}f_{2.}}{f_{.1}f_{12}}$
- Razón de productos cruzados = $RC = \frac{f_{11}f_{22}}{f_{21}f_{12}} = \frac{f_{11} / f_{21}}{f_{12} / f_{22}} = \frac{C_1}{C_2}$

El valor del Riesgo relativo cambia según el tipo de dependencia que tengan las variables, en la forma siguiente:

- El $RR = 1$, informa que no hay asociación entre las variables.
- El $RR > 1$, nos dice que existe asociación positiva.
- El $RR < 1$, indica que existe una asociación negativa.

El caso de la Razón de productos cruzados es distinto, y varía de la siguiente manera:

- El $RC = 1$, implica que la razón de casos que aparece A y \bar{A} , cuando está B presente, y la razón de casos de A y \bar{A} cuando no está presente B , son iguales.
- El $RC < 1$, implica que la razón de casos que aparece A y \bar{A} , cuando se posee el carácter B es menor que la razón de casos de A y \bar{A} cuando no se posee B .
- El $RC > 1$, implica que la razón de casos que aparece A y \bar{A} , cuando está B presente

es mayor que la razón de casos de A y \bar{A} cuando no está presente B .

PR2.2. *Medidas de asociación para tablas rxc.* Estas medidas no tienen en cuenta el signo, que no tiene sentido en este tipo de tablas. Pueden estar basadas o no en el estadístico Chi-cuadrado y las describimos en lo que sigue.

P2.2.1. *Medidas basadas en el estadístico Chi-cuadrado.* Un coeficiente parecido a ϕ es el *coeficiente de contingencia de Pearson*, que se puede utilizar en tablas rxc, y notaremos como C (Ruiz-Maya, Martín, Montero y Uriz, 1995). Toma el siguiente valor:

$$C = \sqrt{\frac{\chi^2}{(\chi^2 + n)}}, \text{ siendo el Max } \{C\} = \sqrt{\frac{\text{Min}\{r-1, c-1\}}{1 + \text{Min}\{r-1, c-1\}}}$$

Cuando C toma el valor 0, ello implica independencia absoluta. Este coeficiente en ocasiones no alcanza el valor 1 (sino que está acotado por el máximo indicado anteriormente), por lo que algunos autores ajustan C dividiéndolo por el máximo posible en una tabla de sus dimensiones. Otro coeficiente es el *de Cramer* (Ruiz-Maya, Martín, Montero y Uriz, 1995), que notamos por V y varía entre 0 y 1 (incluso en tablas no cuadradas). Existen otros como el cuadrado medio del *coeficiente de contingencia* o T de Tschuprow (Ruiz-Maya, Martín, Montero y Uriz, 1995).

$$V = \sqrt{\frac{\chi^2}{n(p-1)}}, \text{ siendo } p = \text{Min}\{r, c\}$$

P2.2.2. *Medidas basadas en la reducción proporcional del error.* Si se quiere predecir un valor de la característica X en un individuo seleccionado aleatoriamente de la población, $P(\text{error regla 1}) = P(\text{Cometer un error en la clasificación}) = \frac{(n - f_i)}{n}$ cuando no se tiene información adicional. Al tener como información adicional el valor de la variable $Y=y_j$ no cogeremos el valor X al azar, más bien tendremos en cuenta (regla 2) cual es el valor más probable de X para $Y=y_j$. Llamaremos medida de la reducción proporcional del error (PRE) al cociente que escribimos a continuación:

$$\text{Medida PRE} = \frac{P(\text{error regla 1}) - P(\text{error regla 2})}{P(\text{error regla 1})}$$

Esta medida (PRE) nos informa sobre el porcentaje de error que se reduce al predecir el valor de la variable dependiente (X), conocido el valor de la variable independiente (Y) (Díaz y Batanero, 2008). Una medida de este tipo bien conocida es la Lambda de Goodman y Kruskal, que se calcula por la expresión siguiente, siendo f_{m+} la frecuencia marginal más grande en filas y f_{mj} la frecuencia marginal mayor en la columna j -ésima.

$$\lambda_x = \frac{(\sum f_{mj}) - f_{m+}}{N - f_{m+}}$$

1.4.4. CONCEPTOS Y PROPIEDADES RELACIONADOS CON EL ANÁLISIS DE TABLAS DE CONTINGENCIA

Como hemos visto, aparecen muchos conceptos y propiedades en los procedimientos ligados a la resolución de los problemas planteados, que podemos clasificar en varios apartados:

Conceptos y propiedades ligados a la clasificación doble o tabla de contingencia

Tanto de la misma tabla de contingencia, como de los problemas y procedimientos descritos para trabajar con ella surgen diferentes conceptos y propiedades que describimos a continuación.

La *variable estadística bidimensional*, o variable que describe el conjunto de pares de valores obtenidos en un experimento. Asociada a ella aparece la *distribución bidimensional* y las frecuencias asociadas, como *frecuencia absoluta* doble (f_{ij}) y la *frecuencia relativa* doble (h_{ij}), tienen la relación siguiente:

$$h_{ij} = f_{ij} / n$$

A partir de la tabla de contingencia (Figura 1.4.1.1), podemos obtener *variables unidimensionales*. Si sumamos por columnas las frecuencias obtenemos como valor de la variable $Y = y_{ij}$, que resulta independiente de la variable X , llamándose *distribución marginal de Y* (Figura 1.4.4.1). Análogamente sumando por filas se obtiene la *distribución marginal de X* (Figura 1.4.4.2). La frecuencia marginal de una cierta variable, es la frecuencia de esa variable sin tener en cuenta ninguna otra.

Figura 1.4.4.1. Distribución marginal de Y

y_1	...	y_j	...	y_c	
f_{1j}	...	f_{ij}	...	f_{cj}	n

Figura 1.4.4.2. Distribución marginal de X Figura 1.4.4.3. Distribución de X condicionada para $Y=y_j$

x_1	$f_{1.}$
x_2	$f_{2.}$
...	...
x_i	$f_{i.}$
...	...
x_r	$f_{r.}$
	n

x_1	f_{1j}
x_2	f_{2j}
...	...
x_i	f_{ij}
...	...
x_r	f_{rj}

Otra distribución unidimensional que se puede obtener, es la *distribución de X condicionada para $Y=y_j$* (Figura 1.4.4.3), extensible a la variable Y condicionada a X. Si tomamos $h(x_i/y_j)$ como la *frecuencia relativa condicional* del valor x_i entre los valores que presenten el carácter y_j se obtiene:

$$h(x_i/y_j) = \frac{f_{ij}}{f_{.j}} = \frac{h_{ij}}{h_{.j}}$$

Análogamente se tiene:

$$h(y_j/x_i) = \frac{f_{ij}}{f_{i.}} = \frac{h_{ij}}{h_{i.}}$$

Como consecuencia, llegamos a la igualdad siguiente:

$$h_{ij} = h(x_i/y_j) h_{.j} = h(y_j/x_i) h_{i.}$$

Conceptos y propiedades ligados al estudio de la asociación y homogeneidad en tablas de contingencia

El concepto más importante que aparece en el análisis de tablas de contingencia es del de *asociación, o dependencia estadística*. Mientras que en una dependencia de tipo funcional a cada valor de una variable X corresponde un solo valor de otra variable Y (dependiente), en el caso de asociación a cada valor de X corresponde una distribución de valores de Y. Además se puede definir una medida de la intensidad de la asociación, que varía entre 0 (independencia total) y 1 (asociación perfecta) utilizando coeficientes como el de correlación (para variables cuantitativas) o diversos coeficientes (para variables cualitativas).

Este término, el de asociación estadística, se usa para hacer referencia a la dependencia entre dos variables de cualquier tipo, mientras que el término correlación se emplea únicamente para variables cuantitativas (Estepa, 1993). En nuestro trabajo sólo analizamos el caso más simple posible: la asociación en variables cualitativas; supondremos además, que los datos están dados en tablas de contingencia, es decir una

tabla que presenta en la primera fila y primera columna los valores de dos variables X e Y , y en las celdas las frecuencias dobles correspondientes a cada par de valores.

En los problemas de juicio de asociación es importante estudiar la *dependencia e independencia* de las variables, lo que puede realizarse a partir de las distribuciones condicionadas. Así que diremos que X e Y son *independientes* cuando se cumpla:

$$h(x_i/y_j) = h_i.$$

Esto significa que la distribución de X no cambia al condicionar por cualquier valor de Y . Otras propiedades de interés son:

- $h_{ij} = h_i \cdot h_j$ para todo i, j
- $h(y_j/x_i) = h_j$ entonces Y no depende de X
- $f_{ij} = \frac{f_i \cdot f_j}{n}$ para calcular las frecuencias teóricas en caso de independencia

Una propiedad interesante de las tablas de contingencia $r \times c$, con $r, c > 2$, es la posibilidad de subdividir la tabla de contingencia en subtablas 2×2 , basándonos en el teorema de la aditividad, haciendo posible calcular la prueba χ^2 para cada una de ellas. Sin embargo, este uso de χ^2 trae el problema de que la suma de variables aleatorias independientes no es exactamente igual al valor del estadístico (Ato y López, 1996).

Conceptos y propiedades ligados al cálculo de probabilidades

Cada una de las variables estadísticas analizadas da origen a una variable aleatoria latente y por tanto la variable estadística doble, que a su vez da origen a la variable aleatoria bidimensional y su distribución. Como hemos visto, en el cálculo de probabilidades, aparecen las probabilidades conjuntas, simples y condicionales. Una vez obtenidas las probabilidades puntuales, conjuntas y condicionadas, es posible la comprobación del teorema de la probabilidad total y teorema de Bayes.

- *Teorema de la probabilidad total.* Considerado un espacio de probabilidad (Ω, Δ, P) y una partición del espacio muestral Ω dada por los sucesos $\{X_1, \dots, X_i, \dots, X_n\}$. Por lo tanto dichos sucesos verifican las dos siguientes propiedades:

$$X_i \cap X_j = 0 \quad \forall i \neq j$$

$$\bigcup_{i=1}^n A_i = \Omega$$

Bajo las hipótesis enunciadas anteriormente se tiene que la probabilidad de

cualquier suceso $Y \in \Delta$ viene dada por

$$\text{Teorema de la probabilidad total} \rightarrow P(Y) = \sum_{i=1}^n P(X_i)P(Y | X_i)$$

- *Teorema de Bayes.* Bajo las mismas condiciones que el Teorema de la probabilidad total, y como consecuencia inmediata de dicho resultado, se verifica

$$\text{Teorema de Bayes} \rightarrow P(X_j / Y) = \frac{P(X_j)P(Y | X_j)}{\sum_{i=1}^n P(X_i)P(Y | X_i)}, \quad \forall j=1, \dots, n$$

También podemos estar interesados en las propiedades de simetría (estando interesados en la distribución conjunta de las dos variables o interesados en la distribución condicional de una de las variables dada la otra) e invarianza (siendo deseable en las medidas de asociación que las variables sean invariantes frente a transformaciones de multiplicar filas y columnas por constantes para poder transformar las distribuciones marginales de la tabla en otras diferentes).

1.4.5. ARGUMENTOS

En el análisis de tablas de contingencia se usan principalmente tres tipos de argumentos:

- Argumentos deductivos, pues todos los procedimientos descritos se basan en deducciones de propiedades o relaciones entre conceptos;
- Argumentos empíricos: pues también las conclusiones de los diversos contrastes utilizan datos obtenidos de experimentos estadísticos.
- Razonamientos de tipo estadístico (contrastados). De hecho el contraste de hipótesis, aunque básicamente es de tipo deductivo, puede considerarse un tipo de razonamiento específico que combina deducción y razonamiento empírico, junto con ciertos criterios de toma de decisión.

1.5. SOFTWARE PARA EL ANÁLISIS DE TABLAS DE CONTINGENCIA

En este apartado hacemos una breve mención al software, que proporciona configuraciones epistémicas localizadas, pues presenta procedimientos, y representaciones característicos (por tanto un lenguaje especial), y requiere utilizar conceptos propiedades y razonamientos (argumentos) que no siempre coinciden con los utilizados sin este recurso. Es mucho el software disponible, pues la mayoría de paquetes estadísticos tienen incorporados una o varias subrutinas para el análisis de

tablas de contingencia. En este punto nos centraremos en el trabajo con algunos softwares usuales con los que se trabaja en la Universidad de Granada en las clases de estadística, es decir SPSS 15.0 y STATGRAPHICS Plus 5.1 y veremos que nos muestran estos dos programas respecto a tablas de contingencia. Existen otros programas de los cuales otros autores tratan como es el BMDP (Aguilera, 2001) y en nuestra experiencia de enseñanza (Capítulo 4) utilizaremos Excel, que no describimos acá, puesto que el programa empleado será analizado con detalle en dicho capítulo.

P1. *SPSS 15.0*. En este programa después de introducir las variables con sus valores podemos trabajar con datos del tipo cualitativos pinchando arriba en el programa Analizar, dentro de él en Estadísticos descriptivos y luego en Tablas de contingencia. Este camino nos abre una ventana para introducir las variables y donde podemos pedirle gran cantidad de cosas como: estadísticos (Chi-cuadrado de Pearson, Chi-cuadrado de la razón de verosimilitud, prueba exacta de Fisher, Chi-cuadrado corregido de Yates, rho de Spearman, coeficiente Phi, V de Cramer, coeficiente de contingencia, lambda y el coeficiente de incertidumbre), gráficos (gráfico de barras agrupadas), frecuencias esperadas, porcentajes, residuos (no tipificados, tipificados y tipificados corregidos) y el contraste Chi.cuadrado. Se puede variar el formato de la tabla (ascendente o descendente) y dar la tabla ya formada o bien introducir un listado de datos para que el programa la forme.

P2. *STATGRAPHICS Plus 5.1*. Con este programa se puede trabajar con datos cualitativos, como es nuestro caso. En la ventana “Descripción”, hay una opción “Datos Cualitativos”, la cuál nos muestra tres opciones distintas para trabajar con datos cualitativos: tabulación, tabulación cruzada y tablas de contingencia.

P2.1. *Tabulación*. Al pinchar en esta opción, nos pide el programa que introduzcamos una variable cualitativa en un lugar llamado “Datos”. En el lado izquierdo nos muestra las variables disponibles de la base de datos. Se nos abre una ventana donde la mitad izquierda contiene las “opciones tabulares”, con un Resumen del Procedimiento y Tabla de Frecuencias; y la mitad derecha contiene las “opciones gráficas”, con Diagrama de Barras y Diagrama de Sectores.

P2.2. *Tabulación Cruzada*. Al pinchar en esta opción, nos pide el programa que introduzcamos las variables cualitativas X e Y , con las que trabajamos, en el hueco “Variable Fila” y la otra en “Variable Columna”. En el lado izquierdo nos muestra las variables disponibles de la base de datos. Se nos abre una ventana donde la mitad izquierda contiene las “opciones tabulares”, con un resumen del procedimiento, tabla de

frecuencias, contraste de Chi-cuadrado y resumen estadístico; y la mitad derecha contiene las “opciones gráficas”, con diagrama de barras, gráfico de mosaico y gráfico tridimensional.

P2.3. *Tablas de Contingencia.* Al pinchar en esta opción, nos pide el programa que introduzcamos las variables cualitativas X e Y , con las que trabajamos, en el hueco “Variable Fila” y la otra en “Variable Columna”. En el lado izquierdo nos muestra las variables disponibles de la base de datos. Se nos abre una ventana donde la mitad izquierda contiene las “opciones tabulares”, con un resumen del análisis, tabla de frecuencias, test Chi-cuadrado y resumen estadístico; y la mitad derecha contiene las “opciones gráficas”, con diagrama de barras, gráfico de mosaico y gráfico tridimensional.

La principal ventaja de cualquiera de estos recursos es la rapidez de cálculo y representación gráfica y la variedad de gráficos y análisis proporcionados. Ello permite emplear mayor porcentaje del tiempo docente en actividades más interpretativas, a la vez que trabajar con datos reales, pues el cálculo ya no es un problema.

1.6. CONCLUSIONES SOBRE EL SIGNIFICADO DE REFERENCIA DE LA TABLA DE CONTINGENCIA

Como resumen del análisis realizado en las anteriores secciones hemos identificado la configuración de objetos y procesos matemáticos que se presenta en la Tabla 1.6.1 y que constituirá el significado de referencia de las tablas de contingencia en nuestro trabajo. Utilizaremos dicho significado para el diseño del proceso de estudio y para la interpretación de los resultados de la evaluación.

Aunque la revisión se limita a un estudio elemental de las tablas de contingencia (pues se podría ampliar, por ejemplo, con el estudio de otros coeficientes o modelos de predicción basados en estas tablas), es ya visible la gran complejidad de este objeto matemático, aparentemente simple y para el cuál se dispone de un tiempo muy limitado de enseñanza, tanto en Psicología como en otras titulaciones. Vemos además, que algunos de los procedimientos que involucra – los contrastes de hipótesis- son a su vez objetos muy complejos que fueron analizados en investigaciones previas (Vallecillos, 1996). Todos estos indicadores apoyan el interés de enfocar una investigación específica sobre la didáctica de las tablas de contingencia.

Tabla 1.6.1. Configuración de objetos y procesos matemáticos

Tipo	Objetos matemáticos
Situaciones -problemas	P1. Resumir información de un conjunto de observaciones bivariantes
	P2. Cálculo de probabilidades asociadas a valores específicos de una o las dos variables.
	P3. Contraste de homogeneidad.
	P4. Contraste de independencia.
	P5. Definir una medida de la intensidad de asociación entre las variables.
Lenguaje	R1. Diagrama de barras adosado.
	R2. Diagrama de barras apilado.
	R3. Gráfico de Mosaico.
	R4. Gráfico tridimensional.
	R5. Lenguaje verbal
	R5. Lenguaje simbólico.
R6. Tablas (de frecuencia y de contingencia).	
Procedimientos	PR1. Contraste de homogeneidad (dos formas alternativa).
	PR2. Contraste para tablas 2x2.
	PR3. Prueba exacta de Fisher.
	PR4. Contraste de independencia.
	PR5. Medidas de asociación para tablas 2x2.
	PR6. Medidas de asociación para tablas rxc (basadas en Chi-cuadrado).
	PR7. Medidas de asociación para tablas rxc (reducción proporcional del error).
	PR8. Lectura e interpretación de la tabla Chi cuadrado.
	PR9. Manejo del software.
Conceptos	C1. Variable simple/valor.
	C2. Variable doble/valor.
	C3. Frecuencia absoluta doble.
	C4. Frecuencia absoluta marginal.
	C5. Frecuencia relativa doble.
	C6. Frecuencia relativa marginal.
	C7. Frecuencia relativa condicional.
	C8. Distribución conjunta.
	C9. Distribución marginal.
	C10. Distribución condicional.
	C12. Probabilidad.
	C13. Casos favorables.
	C14. Casos desfavorables.
	C15. Total/casos posibles.
	C16. Muestra/elemento.
	C17. Porcentaje.
	C18. Razón.
	C19. Proporción.
	C20. Independencia /Asociación.
	C21. Asociación inversa, directa.
C22. Intensidad de la asociación.	
C23. Contraste de hipótesis: Lógica del contraste.	
C24. Hipótesis nula y alternativa.	
C25. Estadístico Chi-cuadrado; distribución Chi-cuadrado, grados de libertad.	
C26. Región crítica y de aceptación.	
Propiedades	PP1. Relaciones entre frecuencias absolutas y relativas.
	PP2. Relaciones entre frecuencias dobles, condicionales y marginales.
	PP3. Igualdad /diferencia de frecuencias relativas condicionales en caso de independencia/dependencia.
	PP4. Igualdad/diferencia de frecuencias relativas condicionales y marginales en caso de independencia/dependencia.
	PP5. Frecuencias esperadas en caso de independencia.
	PP6. Estudio de la asociación implica la comparación de probabilidades.
	PP7. Propiedades de la distribución Chi-cuadrado, grados de libertad.

	PP8. Propiedades del estadístico Chi-cuadrado.
	PP9. Valor p en el contraste Chi-cuadrado. Valor significativo y no significativo.
	PP10. Valor de las medidas de asociación en caso de independencia.
	PP12. Variación de los valores de las medidas de asociación; interpretación.
Arg.	A1. Demostraciones formales algebraicas y/o deductivas.
	A2. Demostraciones informales.
	A3. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar.

1.7.OBJETIVOS DEL TRABAJO

Finalizado la descripción del marco teórico y el análisis epistémico del concepto, estamos en condiciones de fijar los objetivos del trabajo. En el mismo continuamos la línea de investigación iniciada por el grupo “*Teoría de la Educación Matemática y Educación Estadística (Referencia: FQM-126)*” del Plan Andaluz de Investigación de la Universidad de Granada. En particular, se centra en el estudio de la comprensión intuitiva de los estudiantes universitarios de las tablas de contingencia, antes de una enseñanza formal del tema y del aprendizaje adquirido por los mismos como consecuencia de un proceso de estudio específico. Todo ello entendido, según se ha descrito en el marco teórico.

Continuamos también el trabajo de Estepa (1993), donde el énfasis se puso en las concepciones previas y estrategias en el estudio de la asociación, tanto a nivel intuitivo pre-instruccional, como después de una enseñanza de estadística descriptiva basada en ordenadores con estudiantes del curso pre-universitario y futuros profesores. En esta tesis doctoral, además de realizar otro estudio de la comprensión intuitiva de la asociación en las tablas de contingencia, antes de la enseñanza, muy similar al de Estepa (1993) pero con algunas variantes, nos centraremos en la comprensión formal adquirida tras la enseñanza de los contrastes relacionados con las tablas de contingencia y en las medidas de asociación, así como en el cálculo de probabilidades en tablas de contingencia. Por tanto la enseñanza planificada es muy diferente a la de Estepa, incluyendo tanto la parte descriptiva como la inferencia. Otra diferencia es el mayor número de estudiantes y el hecho de que la enseñanza se lleva a cabo dentro de una asignatura reglada obligatoria del plan de estudios, por lo que sus resultados tendrían una aplicación inmediata en la enseñanza del tema en esta titulación.

Para llevar a cabo estos estudios, se requiere una fundamentación previa, que incluye, no sólo el marco teórico descrito en este Capítulo, sino la fijación del significado institucional de referencia del trabajo y el estudio de las investigaciones previas en temas relacionados con nuestro trabajo. De todo ello se deducen los

siguientes objetivos para este trabajo:

- *O1: Realizar un análisis semiótico del objeto matemático “tablas de contingencia”, utilizando los tipos de objetos matemáticos (elementos de significado) considerados en el Enfoque Onto-semiótico*

Dicho análisis se ha llevado a cabo y se describen en este mismo capítulo, donde hemos identificado y descrito los campos de problemas que llevan a la idea de tabla de contingencia, así como los conceptos, propiedades, procedimientos, lenguaje y argumentos involucrados en la resolución de estos problemas. Como consecuencia se ha mostrado la complejidad de este objeto matemático, que involucra ideas de estadística descriptiva e inferencial. Los resultados también servirán para delimitar el significado de referencia del trabajo y posteriormente, diseñar instrumentos de evaluación, las secuencias de enseñanza del tema y para interpretar las respuestas de los estudiantes en las pruebas de evaluación efectuadas.

- *O2: Completar un estado de la cuestión sobre las investigaciones relacionadas con las tablas de contingencia, para situar nuestro trabajo dentro de una línea de investigación más amplia.*

Este objetivo se aborda en el Capítulo 2, partiendo del contenido en la tesis de Estepa (1993) y otros trabajos que sintetizan la investigación en educación estadística, como los de Shaughnessy (1992, 2007); y Shaughnessy, Garfield y Greer (2006). Además del estudio de la asociación se consideran también las investigaciones sobre cálculo de probabilidades en tablas de contingencia, y comprensión del contraste de hipótesis, que no han sido consideradas por el autor citado.

- *O3: Realizar un estudio de evaluación sobre las estrategias intuitivas, juicios de asociación y estimación de la intensidad de la asociación en tablas contingencia.*

Este objetivo se aborda en el Capítulo 3, donde utilizamos un cuestionario con parte de los ítems de Estepa (1993), debidamente modificados en una muestra de 414 estudiantes de tres universidades españolas. Se estudia la percepción de la asociación, estimación del coeficiente, existencia de teorías y previas y nivel de estrategia, relacionando alguna de estas variables entre sí. Se completa un estudio con el análisis semiótico de las estrategias (no realizado por Estepa) y la identificación y clasificación de conflictos semióticos.

- *O4: Diseñar y experimentar un proceso de estudio de las tablas de contingencia dirigido a alumnos de psicología, que contemple las principales dificultades identificadas en el estudio anterior, y en la investigación previa, así como las restricciones curriculares.*

Este objetivo se aborda en el Capítulo 4, donde se analiza un proceso de estudio al que se dedicaron 4 sesiones teóricas y 2 prácticas. Se describe y analiza el material proporcionado al estudiante que consta de cuatro lecciones, ejercicios, actividades prácticas y un programa de cálculo en Excel que facilita los cálculos, junto con su descripción. Se preparó asimismo una página web para poner este material a disposición de los estudiantes. Como resultado del análisis se proporciona el significado institucional pretendido en la enseñanza.

- *O5: Evaluar el aprendizaje logrado por los estudiantes al finalizar el proceso de estudio y comparar cuando sea posible con el conocimiento intuitivo mostrado en el estudio inicial de evaluación.*

Este objetivo se aborda en los Capítulos 5 y 6, donde se describen y analizan los instrumentos de evaluación, presentando el significado evaluado. En el Capítulo 5 se analizan los resultados de evaluación con ítems de verdadero/falso y opción múltiple, estudiando tanto el resultado en cada ítem como las puntuaciones totales y en cada tema, describiendo los puntos sencillos y difíciles en el aprendizaje. También se estudian algunos ítems incluidos en la evaluación final del estudiante.

En el Capítulo 6 se analizan las respuestas de los estudiantes a problema abiertos, describiendo las respuestas observadas en los diferentes apartados los elementos de significado adquiridos y los conflictos semióticos remanentes tras la instrucción. Todos estos resultados permitirán analizar la idoneidad didáctica del proceso de estudio en el Capítulo 7 de conclusiones.

1.8.HIPOTESIS INICIALES

Enunciamos a continuación las hipótesis de nuestra investigación, entendidas como expectativas, y justificando la pertinencia de las mismas.

Como hemos comentado, en esta tesis tenemos previsto realizar dos estudios diferenciados. En el primero de ellos abordamos un estudio de evaluación de los juicios de asociación y estimación de la intensidad de la asociación en tablas de contingencia y

las estrategias intuitivas empleadas en esta tarea. Para esta parte de la investigación, nuestras hipótesis son las siguientes:

- *H1. Se espera que los estudiantes, se vean influenciados por sus teorías previas, afectando de forma significativa en los resultados de juicios de asociación ó estimación de la asociación.*

Esta hipótesis se fundamenta en los resultados de las investigaciones previas, que muestran cómo los sujetos forman teorías sobre la forma en que deben relacionarse algunos acontecimientos y las usan en los juicios de asociación, en lugar de guiarse por los datos. Destacan, en particular las de Chapman (1967), Chapman y Chapman (1969) Murphy y Medin (1985) sobre la correlación ilusoria y en el hecho de que también en la investigación de Estepa (1993) los estudiantes se vieron influenciados por estas teorías.

- *H2. Los estudiantes abordan los juicios de asociación con estrategias intuitivas, la mayoría de las cuáles son incorrectas o parcialmente correctas. El porcentaje de estrategias que alcanzan el nivel 5 en la clasificación de Pérez Echeverría (1990) será pequeño.*

Nos basamos al formular esta hipótesis en los resultados de los estudios que nos anteceden. Por ejemplo, se han encontrado investigaciones de sujetos adultos que basan sus estrategias únicamente en la celda que corresponde a la presencia de los dos caracteres (Smedlund, 1963; Shaklee & Mins, 1982), lo que corresponde al primer nivel de Pérez Echeverría. En otros casos de sujetos adultos comparan las diagonales en frecuencias absolutas, lo que correspondería al nivel 4. En el trabajo de Pérez Echeverría (1990) aparecen sujetos en todos los niveles descritos por la autora siendo el porcentaje de sujetos en el nivel 5 únicamente del 12,4%; por otro lado en el estudio de Estepa (1993) en ninguno de los problemas que plantea el porcentaje de sujetos este nivel alcanzaba el 5% .

- *H3. Las variables de tarea utilizadas en el estudio inicial de evaluación así como la estrategia utilizada, influyen en el juicio de asociación y en la estimación de la intensidad.*

El cuestionario utilizado en dicho estudio es una adaptación de cuatro de los ítems que se utilizó en las investigaciones de Estepa (Estepa ,1993; Estepa y Batanero, 1995), habiéndose respetado las variables consideradas por el autor, quien encontró que influían en la estimación de la asociación y el juicio de asociación. El autor no estudió

la influencia del tipo de estrategia (correcta, parcialmente correcta o incorrecta) sobre estos juicios. Pretendemos completar en este punto la investigación de Estepa.

- *H4. Como resultado del análisis semiótico de las estrategias de los estudiantes en el estudio inicial de evaluación se espera describir y clasificar diferentes conflictos semióticos que expliquen las estrategias incorrectas o parcialmente correctas en los juicios de asociación.*

En este trabajo se utilizan algunas nociones teóricas relacionadas con el Enfoque Ontosemiótico desarrollado por Godino y su equipo de colaboradores (Godino, 2002b; Godino, Batanero y Font, 2007). Pretendemos realizar un análisis semiótico de las estrategias de los estudiantes apoyados en este marco. Esperamos que este análisis nos ayudará al reconocimiento de conflictos semióticos, que permiten explicar las dificultades de los estudiantes, al igual que ha ocurrido en otras investigaciones (Mayén, Díaz y Batanero, 2009; Arteaga, 2011; Contreras, 2011).

Por otro lado, tenemos previsto realizar y evaluar una experiencia de enseñanza relacionada con las tablas de contingencia, lo que nos permitira observar el aprendizaje formal adquirido por los estudiantes. Para esta parte de la investigación planteamos las siguientes hipótesis:

- *H5. Se espera que los resultados de las pruebas de evaluación indiquen la adquisición de la mayoría de elementos de significado incluidos en el proceso de estudio por un porcentaje de alumnos superior al 50%.*

La construcción de cuestionarios para la evaluación tendrá en consideración todos los elementos de significado, para la observación de la adquisición de los mismos. No tenemos antecedentes que apoyen esta hipótesis, pues no los hemos encontrado sobre enseñanza de las tablas de contingencia.

- *H6. Algunas concepciones iniciales erróneas descritas por Estepa (1993) serán superadas por los estudiantes participantes al finalizar la enseñanza.*

La enseñanza planificada tendrá en cuenta las dificultades descritas en la investigación previa, así como las observadas en el estudio de evaluación inicial, por lo que esperamos contribuya a superarlas. Por otro lado, algunas investigaciones que describen experimentos de enseñanza de la estadística en Psicología, apoyados como el nuestro en el ordenador para la realización de los cálculos muestran una mejora de la

comprensión en los estudiantes. (Díaz, 2007). También en la enseñanza descrita por Estepa (1993) hubo alguna mejora, que esperamos sea mayor en nuestro experimento.

- *H7. Esperamos que las estrategias empleadas en los juicios de asociación después de la enseñanza sean en su mayoría correctas; ampliando el número de estrategias formales. adquiridas en la instrucción y desapareciendo prácticamente las estrategias de niveles 4 o inferior en la clasificación de Pérez Echeverría.*

Como mejora por parte de nuestra enseñanza, pensamos que las estrategias planteadas por los estudiantes serán de mayor formalidad y válidas para obtener resultados correctos. También esperamos se abandonen las estrategias basadas en el uso de una sola celda o de una sola distribución.

- *H8. Los estudiantes mostrarán también competencia en la realización de contrastes de homogeneidad y en la interpretación de diferentes coeficientes de asociación en problemas abiertos.*

La complejidad de los contrastes de hipótesis se ha descrito en múltiples investigaciones (por ejemplo Vallecillos, 1996; Vallecillos y Batanero, 1997), pero casi todas ellas se han centrado en los contrastes paramétricos. Parte de estas dificultades podrían aparecer en los contrastes de independencia y homogeneidad, realizados a partir de tablas de contingencia. No obstante, esperamos que la enseñanza realizada contribuya a que una proporción alta de estudiantes adquiera la suficiente competencia para llevarlos a cabo.

- *H9. Se espera mostrar una adecuada idoneidad del proceso de estudio diseñado.*

Utilizaremos indicadores de idoneidad didáctica de nuestro marco teórico (Godino, 2011) para analizar el proceso de estudio, esperando que la valoración final del mismo sea positiva en la mayor parte de los componentes de dicha idoneidad (epistémica, cognitiva, afectiva, interaccional, mediacional). Esta valoración será también útil para plantear la revisión futura de nuestra enseñanza y superar los puntos débiles encontrados.

1.9. ESTRUCTURA DEL TRABAJO Y CARACTERÍSTICAS METODOLÓGICAS GENERALES

En esta investigación se utiliza una metodología mixta entre los métodos cuantitativos y cualitativos, ya que analizamos variables de ambos tipos que utilizamos en la investigación. Por otro lado, la parte empírica de la investigación, la dividiremos en dos partes, en cada una de los cuales se ha empleado diferentes metodologías, tanto para la recopilación de la información, como en el análisis de datos.

Como en todo estudio en que se trata de evaluar la comprensión (en este caso de las tablas de contingencia) hay que tener en cuenta que trabajamos con constructos inobservables (León y Montero, 2002), por lo que sus características deben ser inferidas de las respuestas de los alumnos. Con los instrumentos de medición utilizados en el estudio tratan de inferir conocimientos y capacidades de los sujetos, que no son accesibles por simple observación o encuesta (Castro Posada, 2001; Barbero, 2003).

En la Tabla 1.9.1 se resumen la metodología utilizada. Estas partes se describen a continuación.

Tabla 1.9.1. Métodos empleados en las diferentes partes que componen el trabajo

Partes	Descripción	Muestra	Material	Análisis	Resultados
Parte 1	Estudio de las concepciones intuitivas antes de la enseñanza	414 estudiantes de psicología de 3 universidades andaluzas	Cuestionario de cuatro ítems, adaptado de Estepa (1993)	Análisis descriptivo; test Chi-cuadrado; Análisis de varianza multivariante, pruebas a posteriori; Análisis semiótico	Intensidad de la asociación; Juicios de asociación; Estrategias en los juicios de asociación; Clasificación y comparación de estrategias; Conflictos semióticos
Parte 2	Diseño y estudio de una experiencia de aula	104 estudiantes de psicología de la universidad de Granada	Material de enseñanza (apuntes, pagina web, hoja Excel...), Instrumentos de evaluación: cuestionarios y problemas abiertos	Análisis descriptivo; Análisis de ítems: frecuencias, discriminación, dificultad; Analisis a priori del significado; Análisis de fiabilidad; Análisis cualitativo de respuestas a problemas abiertos	Resultados por ítem; Identificación de puntos sencillos y difíciles en el aprendizaje; Puntuaciones totales y parciales; Clasificación de respuestas en cada apartado de los problemas abiertos; Descripción de errores y conflictos semióticos

Parte 1. Estudio inicial de evaluación,

La primera parte será presentada en el Capítulo 3, donde se realiza un estudio de evaluación a una muestra de 414 estudiantes del primer curso de psicología mediante un cuestionario formado por ítems adaptados de Estepa (1993). Se trata de una investigación *explicativa* pues nuestra preocupación se centra en determinar los orígenes o las causas de un determinado conjunto de fenómenos (estrategias y juicios de asociación). El objetivo es conocer por qué suceden ciertas respuestas, a través de las condiciones en las que se producen (Losada y López Feal, 2003). Puesto que podemos intervenir (mediante el diseño del cuestionario) en la manipulación de las variables independientes se trataría de un estudio cuasi-experimental (pues las muestras no son aleatorias), más concretamente un cuasi-experimento, pues no hay asignación aleatoria (Cook y Campbell, 1979).

En estos ítems se contemplan las siguientes variables independientes: tipo de asociación, intensidad, coincidencia de los datos con la teoría previa y explicación de la asociación. Otra variable independiente es la universidad de procedencia del estudiante. Las variables dependientes consideradas en el estudio son los juicios de asociación, estrategias utilizadas, número de casillas utilizadas en la estrategia (que permite clasificarlas según el nivel descrito por Perez Echeverria (1990), y estimación de la asociación.

Sobre estas variables se realizaron diferentes tipos de análisis, pues se trata de un estudio que combina la metodología cuantitativa (estudio de la precisión en la estimación de la asociación y los juicios de asociación) y el enfoque cualitativo (en el análisis de estrategias).

- Como métodos estadísticos se utilizan estadísticos descriptivos, contraste Chi-cuadrado, para comprobar la homogeneidad de respuestas según universidad de procedencia; análisis de varianza multivariante, y contrastes pos- hoc de medias, siguiendo los procedimientos de Tukey (1953) y Scheffé (1953).
- El análisis semiótico de las respuestas a estos problemas permite describir el conocimiento matemático de los psicólogos futuros, definiendo categorías de respuestas correctas e incorrectas y conflictos semióticos.

Parte 2. Diseño y estudio de una experiencia de aula, relacionada con las tablas de contingencia,

La segunda parte será presentada en los Capítulos 4, 5 y 6. Se diseña una

experiencia de aula (Capítulo 4) en un periodo de cuatro horas teóricas, dos horas prácticas y dos horas de realización de cuestionarios. El diseño de una propuesta concreta de enseñanza para unos estudiantes particulares en contexto dado entra dentro de los estudios curriculares, que debe tener en cuenta cuatro grandes cuestiones: ¿Qué enseñar?, ¿cuándo enseñar? ¿cómo enseñar?, y ¿qué, cómo y cuándo evaluar? (Bolívar, 1998).

Se trata, en consecuencia de un experimento, puesto que el experimentador produce un evento, puede describir sus condiciones para que otros experimentadores puedan reproducirla (Losada y López Feal, 2003); más concretamente un cuasi-experimento, pues no hay asignación aleatoria (Cook y Campbell, 1979). Además se tuvo en cuenta la dimensión instruccional de nuestro marco teórico (Godino, 1996), y en particular, el concepto de trayectoria didáctica. Los principios instruccionales propuestos por Cobb y McClain (2004) y Garfield y Ben-Zvi (2008), también se han tenido en consideración.

En esta actividad participaron 104 estudiantes de la Universidad de Granada, que anteriormente habían colaborado en la parte 1 descrita anteriormente. Debido a que la muestra es de tamaño moderado e intencional, la finalidad del análisis será exploratoria, en cuanto a la intención de generalización a poblaciones diferentes de estudiantes.

Se comienza con un periodo de elaboración de materiales didácticos (descritos en el Capítulo 4) para su uso durante la experiencia, a los cuales se aplica un análisis de contenido para asegurar su validez. Estos materiales están a disposición de los estudiantes durante el desarrollo de la enseñanza.

Como instrumento de evaluación se crearon cuestionarios formados por ítems de verdadero/falso y opción múltiple, así como problemas abiertos; los datos recogidos por los mismos se analizan en los Capítulos 5 y 6. En su construcción se siguieron normas psicométricas para asegurar su calidad de los instrumentos (Osterlind, 1989; Thorndike, 1989; Martínez, 1995; Martínez, Hernández, Hernández, 2006). Se estudió la validez de contenido de los instrumentos, a partir del análisis a priori del significado evaluado. También se hace un estudio de los indicadores psicométricos del cuestionario (dificultad, discriminación) (Thorndike, 1989; Martínez, 1995), y se estudia su fiabilidad, mediante el coeficiente de consistencia interna de Cronbach (Carmines y Zeller, 1979; Meliá, 2001).

Se utilizó durante el periodo de enseñanza una metodología observacional, cuyo procedimiento *según los sujetos*, es “de barrido”, pues se hace un registro de la conducta

de un grupo (los alumnos) durante intervalos regulares (las sesiones de clase); *según el tiempo*, “continuo”, donde se anotan todos los casos en que se da una categoría de conducta (las preguntas que hacen los alumnos) (Anguera, 1990; Anguera, 1993a y 1993b). Las *técnicas de registro* de la observación fueron escalas de evaluación, donde un observador rellena un cuestionario sobre la conducta que observa y como, *instrumentos de registro* en nuestro caso se utilizaron fichas de observación de la clase por una tercera persona (Anexo 4) y grabaciones de audio, para hacer un informe completo de lo ocurrido en clase. Finalmente, añadir que hemos tenido en consideración los requisitos de la observación de fiabilidad y validez (Anguera, 1990; Anguera, 1993a y 1993b; Losada, 1999).

El estudio de evaluación realizado después de la enseñanza se considera un estudio exploratorio (Danhke, 1989; Hernández, Fernández, Baptista, 1994), ya que el objetivo fue examinar un tema o problema de investigación poco estudiado.

Respecto a los métodos de análisis se utilizan como métodos estadísticos en esta fase: estadísticos descriptivos, índices de dificultad y discriminación. Finalmente se realiza una relación entre respuestas al cuestionario y en la prueba final, que nos proporcione si se refleja un efecto de olvido.

Los datos recogidos con los problemas abiertos, se analizaron cualitativamente, clasificando las diferentes respuestas en cada apartado, mediante un proceso inductivo y cíclico, hasta llega a un número suficiente de categorías para mostrar la diversidad de soluciones. Se aplicó un análisis de contenido, que nos permite pasar del texto a una serie de variables estadísticas de las cuáles posteriormente se realizan inferencias (Ghiglione y Matalón, 1991). Se realiza un estudio descriptivo de los datos obtenidos en cada problema y apartado y un estudio global de la corrección en el conjunto de problemas.

Por último, queremos que conste que, tras este diseño de investigación, esperamos obtener unos resultados de interés para la enseñanza de las tablas de contingencia a estudiantes de Psicología.

CAPÍTULO 2.

JUICIOS DE ASOCIACIÓN Y CÁLCULO DE PROBABILIDADES EN TABLAS DE CONTINGENCIA

2.1. Introducción
2.2. Asociación, condicionamiento y causación
2.2.1. Introducción
2.2.2. Elementos de la relación causal
2.2.3. Teorías filosóficas sobre causalidad
2.2.4. Causalidad desde el punto de vista estadístico
2.3. Juicios de asociación
2.3.1. Desarrollo evolutivo del concepto de asociación
2.3.2. Estrategias en los juicios de asociación
2.3.2.1. Investigaciones en psicología
2.3.2.2. Sesgos en la detección de la asociación y efecto de las teorías previas
2.3.2.3. Investigación en didáctica de la matemática
2.3.3. Concepciones sobre la asociación
2.3.4. Evolución después de la enseñanza
2.3.4.1. Evolución de estrategias en juicios de asociación
2.3.4.2. Juicios de asociación usando ordenador
2.3.4.3. Evolución de las concepciones
2.3.4.4. Actos de comprensión de la asociación
2.4. Calculo de probabilidades a partir de tablas de contingencia
2.4.1. Investigaciones con estudiantes de psicología
2.4.2. Trabajos con futuros profesores
2.4.3. Otros trabajos
2.5. Comprensión del contraste de hipótesis
2.6. Conclusiones del estado de la cuestión

2.1.INTRODUCCIÓN

Batanero, Godino, Green, Holmes y Vallecillos (1994) indican que una gran parte de la investigación didáctica trata de caracterizar los errores y dificultades de los estudiantes sobre los principales conceptos estadísticos elementales: entre ellos incluyen los estudios sobre juicios de asociación en tablas de contingencia, tema en el que nos centramos en esta investigación.

Con objeto de fundamentar nuestro trabajo, en este capítulo hacemos una síntesis de las investigaciones previas relacionadas con las tablas de contingencia, que podemos clasificar en dos grupos principales: (a) Estudios de las estrategias, sesgos y concepciones en relación con los juicios de asociación en tablas de contingencia; y (b) Estrategias y sesgos en el cálculo de probabilidades a partir de tablas de contingencia.

Asimismo se incluye un análisis de algunas investigaciones que describen dificultades y errores asociados al contraste estadístico de hipótesis, pues en el estudio formal de las tablas de contingencia los estudiantes han de estudiar los contrastes de asociación e independencia, basados en la distribución Chi-cuadrado.

Para revisar estas investigaciones, partimos del trabajo de Batanero et al. (1994), así como los de Pérez Echeverría (1990) y Estepa (1993, 2004), estos últimos centrados en el tema de asociación estadística. También hemos revisado otros trabajos posteriores de síntesis, como los de Shaughnessy (1992, 2007); y Shaughnessy, Garfield y Greer (2006) y Engel y Sedlmeier (2011), se analizaron también las investigaciones relacionadas publicadas en los congresos ICOTS (International Conference on Teaching Statistics) y las revistas *Statistics Education Research Journal*, *Journal of Statistics Education*, así como las principales revistas y congresos de educación matemática.

En lo que sigue describimos estos tres tipos de investigaciones, incluyendo primero un análisis de las relaciones entre los conceptos de asociación, condicionamiento y causación, puesto que la finalidad de realizar un juicio de asociación es con frecuencia la búsqueda de relaciones causales; pero, sin embargo, la existencia de asociación no siempre se debe a la causalidad.

2.2. ASOCIACIÓN, CONDICIONAMIENTO Y CAUSACIÓN

2.2.1. INTRODUCCIÓN

En el Capítulo 1 se realizó un análisis del significado institucional del objeto matemático “tablas de contingencia”, observando que uno de los conceptos claves ligado a las mismas es la asociación estadística, que extiende la idea de dependencia funcional, y se utiliza en muchos métodos de gran variedad de ciencias (Batanero, 2001). Mientras que en una dependencia de tipo funcional a cada valor de una variable X corresponde un solo valor de otra variable Y (dependiente), en el caso de asociación a cada valor de X corresponde una distribución de valores de Y . Además se puede definir una medida de la intensidad de la asociación, que varía entre 0 (independencia total) y 1 (asociación perfecta) utilizando coeficientes como el de correlación (para variables cuantitativas) o el coeficiente de contingencia C (para variables cualitativas).

Este término, el de asociación estadística, se usa para hacer referencia a la dependencia entre dos variables de cualquier tipo, mientras que el término correlación se emplea únicamente para variables cuantitativas (Estepa, 1993). En nuestro trabajo

sólo analizamos el caso más simple posible: la asociación en variables cualitativas; supondremos además, que los datos están dados en tablas de contingencia, es decir una tabla que presenta en la primera fila y primera columna los valores de dos variables X e Y , y en las celdas las frecuencias dobles correspondientes a cada par de valores.

El estudio de la asociación es un primer paso para describir relaciones de causalidad. Sin embargo, si dos sucesos covarían ello no implica que el primero (A) sea causa del segundo (B). La covariación entre los acontecimientos A y B puede deberse a que, efectivamente, el antecedente A sea la causa del consecuente B . Esta relación causal entre A y B puede existir, aunque la consecuencia se manifieste en primer lugar. Por ejemplo, la aparición de un síntoma (consecuencia) es previa a la constatación de la enfermedad (la causa) que lo provoca (Perales, Catena, Ramos y Maldonado, 1999).

Por otro lado puede existir la asociación sin que se de la causalidad. En general, la relación de causalidad es claramente diferente de la de asociación y su diferencia entre ellas es un tema central para la filosofía de la causalidad. En lo que sigue se analizan las relaciones entre asociación y causalidad.

2.2.2. ELEMENTOS DE LA RELACION CAUSAL

Una cuestión importante en todas las ciencias es determinar hasta qué punto cuando dos fenómenos ocurren conjuntamente (asociación) uno de ellos produce el otro (causalidad). Por ejemplo, Álvarez y Pérez (2004) estudian la causalidad en el campo de la medicina, indicando que:

“El concepto de causalidad es parte de la vida común, pues el hombre desde siempre ha buscado el por qué de las cosas, como una forma de entenderse y adaptarse al mundo” (p. 467).

Dichos autores explican que uno de los objetivos de la medicina es la identificación de los factores o agentes que causan las enfermedades, indicando la importancia de la causa para establecer tratamientos, y aplicar medidas preventivas. Entre los tipos de causa diferencia entre causa única o múltiple, siendo habitual la causa única en las enfermedades infecciosas, mientras que en las enfermedades crónico-degenerativas es más común la existencia de múltiples causas.

La relación causal ha sido definida en forma diversa. Por ejemplo Bradford-Hill, (1992) define la causa como un evento o un estado de la naturaleza que inicia o permite otra secuencia de eventos que resultan en un efecto. Clark (1994) indica que la causa es lo que se considera como fundamento u origen de algo, su razón, motivo u origen, o el

factor que es posible o conveniente alterar para producir, modificar o prevenir un efecto.

Álvarez y Pérez (2004) identifican los siguientes elementos de la relación causal: un elemento inicial, un elemento final y la relación entre ambos. El elemento inicial varía según acciones intencionales, no intencionales, atributos o constructos. El elemento final es el cambio respecto a su situación previa, es decir, respecto a lo que hubiera ocurrido en ausencia de causa. La relación entre el elemento inicial y el elemento final es una función o condición, que puede ser de varios tipos:

- Necesaria y suficiente: si A , entonces B y si B , entonces A .
- Necesaria y no suficiente: si A , entonces B o no B y si B , entonces A .
- No necesaria y suficiente: si A , entonces B y si B , entonces A o no A .
- No necesaria y no suficiente: si A , entonces B o no B y si B , entonces A o no A .

2.2.3. TEORÍAS FILOSÓFICAS SOBRE CAUSALIDAD

Como se ha indicado, la diferencia entre ambos conceptos es un tema central para la filosofía de la causalidad. En lo que sigue realizamos una síntesis de las teorías más conocidas, que se encuentran resumidas en Cook y Campbell (1979) y Estepa (1993). Al igual que dicho autor, no realizamos un estudio completo de las teorías causales, por no ser objetivo de la investigación. Tan solo presentamos una síntesis de las siguientes teorías filosóficas: la tradición positivista, teoría esencialista, teoría de John Stuart Mill, teoría de Popper, la teoría de la causación activa y la perspectiva crítica-realista.

Primeros pasos

Álvarez y Pérez (2004) mencionan a Tales de Mileto, como autor de una de las primeras teorías causales, en su interés por buscar las causas más profundas de las cosas; dividió éstas en causas inmediatas y esenciales o últimas. Pozo (1987) indica que la primera teoría sistemática sobre este tema es debida a Aristóteles, quien en su trabajo diferenció la existencia de cuatro tipos de causas diferentes, que son: la causa material (soporte material), la causa formal (cualidad de la cosa), la causa eficiente (el responsable del efecto) y la causa final (lugar al que tiende la cosa), las cuales fueron aceptadas por la comunidad científica hasta el tiempo de Galileo, momento en el cual este reduce estos cuatro tipos de causa a la causa eficiente, es decir la que produce un efecto.

Teoría positivista

Para David Hume, la causalidad es la forma con la que se expresan las leyes científicas basadas en una generalización de la experiencia y con capacidad predictiva. Hume pensaba que las observaciones de los fenómenos naturales eran una base débil para las inferencias causales y que el descubrimiento de sus causas requería un proceso intelectual que trascendiese las impresiones de nuestros sentidos. La tradición positivista, representada por Hume, exige, para hablarse de causa, la contigüidad entre el efecto presumido y la causa, la precedencia temporal de la causa sobre el efecto (la causa antecede al efecto), y que la causa ha de presentarse siempre que se dé el efecto (Rivadulla, 1991).

Para Pozo (1987) los principios causales de la tradición positivista son: determinismo causal, constancia, condicionalidad, asimetría y productividad. En esta tradición aparece Russell, quien sugiere que en algunas ciencias, tales como astronomía o Física, en que se daban estos supuestos, podría prescindirse del concepto de causa, ya que se obtenían buenas predicciones a partir de relaciones funcionales.

Cook y Campbell (1979) al estudiar los procedimientos de investigación de las ciencias sociales, indican que el concepto de causa es más importante en estos campos, en comparación con las ciencias físicas o naturales, donde es más sencillo aplicar los siguientes procedimientos:

- Aislamiento experimental, llevando a cabo ensayos en laboratorio con condiciones muy controladas; sin embargo en las ciencias sociales es difícil controlar todas las variables.
- Usualmente las variables dependientes analizadas en las ciencias naturales no varían a lo largo del tiempo si no se produce un tratamiento, o varían poco (por ejemplo, el peso de un adulto), mientras que en las sociales variables como actitudes, sentimientos, etc., varían en el sujeto más rápidamente.
- En las ciencias naturales se dispone de teorías explícitas y precisas que especifican el tamaño esperado de los efectos, y se dispone de instrumentos de medición suficientemente precisos para medir el efecto esperado. Los instrumentos de medición son mucho más difíciles de construir en las ciencias sociales.

El positivismo está en conflicto con la idea dualista de evidencia y realidad, idea reflejada en el trabajo de Carmines y Zeller (1979). Además aparece en el positivismo el dogma de emplear definiciones operativas para los términos teóricos, que llevo al uso repetido de instrumentos particulares de medida, concepción abandonada por inaceptable en las ciencias sociales, ya que las medidas implican muchas variables teóricas conocidas, desconocidas y muchas hipótesis no probadas (Estepa, 1993).

Otro representante de esta escuela según Pozo (1987) es Kant, para quien todo conocimiento comienza con la experiencia, siendo la causalidad un a priori necesario para el conocimiento racional, y señalando que todo cambio tiene que tener una causa.

Teoría esencialista

En esta escuela, el término causa se utiliza para analizar la relación entre variables que explican fenómenos en forma necesaria y suficiente. Por tanto no se consideran variables que a veces producen el fenómeno pero no siempre (Cook y Campbell, 1979). Además en esta teoría no se requiere que la causa preceda al efecto, pero sí se requiere que las dos variables estén simultáneamente relacionadas, pues si no, podrían ocurrir sucesos intermedios que influyesen en el efecto.

Russell (1973) comenta en su trabajo una cuestión sobre este tema, consistente en como se podría especificar la longitud de tiempo entre el cambio de *A* y la respuesta de *B*. Pues se dan situaciones en que al ser la respuesta de *B* simultánea con *A*, podría ocurrir que una medida retardada de *B* pudiese fallar en detectar esta respuesta, o el caso de que una respuesta sea retardada, una medida prematura de *B* no la detectaría. En consecuencia no se sabe si es o no necesario incluir el intervalo de tiempo fijo para establecer la relación causal.

Según Estepa (1993), el uso de relaciones causales infalibles y necesarias, por parte de los esencialistas, lleva a la confusión de una alta correlación con causa y a la duda de por qué las condiciones necesarias y suficientes han de preceder al efecto. Las relaciones funcionales llevan a causación reversible y las causas bi-direccionales ocasionan problemas en los estudios correlacionales, pues no se distingue cual es la causa y cual el efecto. En los estudios experimentales en ocasiones no se controla una variable que afecta a la dependiente, o se puede manipular un constructo causal en un estudio y luego, manipular el efecto producido para ser introducido como causa en un segundo estudio.

Teoría de John Stuart Mill

Rivadulla (1991) indica que Mill mantenía que la inferencia causal depende de tres factores:

- La causa ha de preceder al efecto.
- La causa y el efecto han de estar relacionados.
- Han de eliminarse otras explicaciones causa-efecto sobre el fenómeno.

Con estas ideas trabaja el método de concordancia, el método de diferencias y el método de variación concomitante. El método de concordancia consiste en comprobar que el efecto se presenta cuando está presente la causa, el método de diferencias indica que el efecto ha de estar ausente cuando está ausente la causa, y el método de variación concomitante llega a la conclusión que cuando se observan causa y efecto, la inferencia causal tendrá mayor fuerza cuando se puedan descartar otras interpretaciones de la covariación entre las variables.

Mill constata que al comparar casos donde una explicación alternativa particular a la inferencia válida esté y no esté operando proporciona la clave para comprobar si dicha explicación alternativa puede o no explicar la relación observada.

Popper y el contraste de hipótesis

Otro autor de gran influencia en la estadística ha sido Popper (1962), quien se preocupa por dar una solución al problema planteado por Hume sobre la validez de la inducción, es decir, si un conocimiento establecido en forma inductiva tiene validez lógica o no. Popper trata de analizar si es posible dar una validez probabilística a dicho problema, y muchos estadísticos estudiaron posteriormente este problema, sin lograr una solución satisfactoria. Según Messick (1992), este resultado lleva a negar la posibilidad de conocimiento confirmatorio basado en la generalización de observaciones particulares a proposiciones científicas generales. Tanto Popper como Messick están de acuerdo en que las teorías científicas pueden usarse para generar predicciones cuantitativas respecto a los resultados obtenidos de un experimento científico, aunque también indican que los datos nunca reproducen exactamente una predicción cuantitativa. Ello produce, como consecuencia fallos al confirmar o falsar una teoría, que son descritos por los errores admitidos en el contraste de hipótesis.

Según Messick (1992), Popper enfatiza la ambigüedad de la confirmación de

una hipótesis de modo inductivo, ya que dicha aceptación sólo demuestra que la teoría resulto contrastada, y que no ha podido ser rechazada. Sin embargo, sólo unos pocos casos que la contradigan pueden llevar a rechazarla. La asimetría entre el rechazo de una hipótesis y la confirmación inconclusa es la clave del énfasis de Popper en el concepto de “falsación”.

Según Cook y Campbell (1979), el punto de vista de Popper se apoya en aceptar que las teorías pueden ser comparadas entre sí. Pero este punto de vista es rechazado por otros filósofos como Kuhn o Feyerabend, quienes defienden que las observaciones están impregnadas de teoría bajo la que son recogidos y por tanto no pueden ser comparables entre sí. Por otro lado Cook y Campbell mencionan el hecho de que algunos filósofos de la ciencia exageran el peso del contraste de hipótesis en el avance científico, pues en ciertos casos la experimentación exploratoria no guiada por teorías formales ha sido el fruto de grandes descubrimientos científicos (Estepa, 1993).

La teoría de la causación activa

Collingwood y Suppes consideran que causa implica manipulación, considerando que la manipulación de la causa conlleva al efecto, aunque la probabilidad de obtener el efecto disminuye cuanto más “social” sea el fenómeno y cuanto más tiempo pase entre causa y efecto (Cook y Campbell, 1979). Collingwood distingue entre tres sentidos del término causa:

- El primer sentido, donde la causa es un acto libre y deliberado de un ser consciente y esto es lo que le motiva a actuar de una manera determinada.
- Un segundo sentido de causa es un suceso natural que produce o impide que ocurra algo.
- Como tercer sentido la causa explicaría un fenómeno de la forma más completa posible de modo que llegásemos a la causa inevitable de un suceso único (Estepa, 1993).

Teoría crítica-realista

Según Estepa (1993), esta perspectiva de la causalidad asume que la disposición humana para inferir relaciones causales es fruto de la evolución biológica de los procesos mentales-cerebrales que ha resultado en una unidad psíquica referida a la causación. La unidad es adaptiva por el hecho de que en ocasiones detecta procesos causales ocurridos en el mundo real, pero sin ajustarse totalmente en su percepción a lo

que ocurre en la realidad. En esta visión se reconocen las percepciones causales como “subjetivas”, construidas o realizadas por la mente, remarcando que muchas son verdades sobre la naturaleza del mundo, con un contenido que puede ser objetivamente falso o cierto.

Esta teoría es realista por asumir que la relación causal existe en el exterior de la mente humana, siendo además realista-crítica por asumir no poder ser percibida en su totalidad, por el hecho de ser limitados nuestros sentidos. Iniciada desde tiempos de Kant, otros autores como Popper o Bunge han sido partidarios de la misma.

La ciencia actual

Álvarez y Pérez (2004) sugieren que la mayoría de los filósofos han pensado que la noción de causalidad es indispensable para la ciencia, aunque esta posición ha sido cuestionada por Russell por considerarla circular. Cuando una ciencia alcanza una cierta experiencia, la noción de causa desaparece y el principio de causalidad es un concepto que pertenece a las distintas imágenes del mundo, y a la interpretación de las formulas científicas contemporáneas. Los autores analizan varios modelos causales en biología:

El Modelo de Koch-Henle (Fletcher, Fletcher y Wagner, 1998) fue sugerido en el estudio de la tuberculosis y las enfermedades infecciosas y es el resultado de un conjunto de reglas de decisión, que son:

- El microorganismo (causa) debe encontrarse siempre en los casos de enfermedad (efecto);
- El microorganismo deberá poder ser aislado en cultivo, demostrando ser una estructura viva y distinta de otras que puedan encontrarse en otras enfermedades. Es decir, la causa ha de ser aislada de otras posibles;
- El microorganismo cultivado deberá ser capaz de producir la enfermedad en el animal de experimentación (la causa produce el efecto);
- Es preciso recuperar el microorganismo a partir del animal e identificarlo (se debe constatar la causa).

El modelo de Bradford-Hil (1992), resulta paradigmático en el estudio de las enfermedades no infecciosas, llegando a una relación causal no necesaria y no suficiente, donde aparecen los siguientes criterios:

- Fuerza de asociación, medida con los índices estadísticos apropiados.

- Congruencia entre distintos observadores, en diferentes lugares, tiempos y circunstancias.
- Especificidad de las causas.
- Temporalidad.
- Gradiente biológico en la relación dosis-respuesta.
- Plausibilidad biológica.
- Coherencia con otros conocimientos.
- Evidencia experimental.
- Analogía con otras relaciones causales.

El modelo de Rothman (1995) en el campo de la epidemiología muestra las relaciones multicausales, definiendo como causa a todo acontecimiento, condición o característica que juega un papel esencial en producir un efecto, caracterizándose principalmente por:

- Ninguna de las causas componentes es superflua.
- No exige especificidad, pues un mismo efecto puede ser producido por distintas causas suficientes.
- Una causa componente puede formar parte de más de una causa suficiente para el mismo efecto. Si una causa componente forma parte de todas las causas suficientes de un efecto se le denomina causa necesaria.
- Una misma causa componente puede formar parte de distintas causas suficientes de distintos efectos.
- Dos causas componentes de una causa suficiente se considera que tienen una interacción biológica, es decir, ninguna actúa por su cuenta. El grado de interacción puede depender de otras causas componentes.

Son numerosos los ejemplos de trabajos de otras especialidades donde se aborda el término causalidad. Algunos ejemplos en la física los proporcionan la investigación de González (2005) donde uno de los dos objetivos es mostrar cómo la interpretación de Born no supone abandonar una explicación causal de los fenómenos físicos. En Retolaza (2007), se observa el problema de la causalidad en las ciencias económicas. En la medicina Laborda (2007) estudia las diferentes teorías sobre el nexo de causalidad, orientado, a la actividad médica.

2.2.4. CAUSALIDAD DESDE EL PUNTO DE VISTA ESTADÍSTICO

Como hemos visto en el apartado anterior, numerosos filósofos y científicos han tratado de relacionar “causalidad” y asociación, considerando que si una variable A es “causa” de otra variable B entonces A está correlacionada positivamente con B , o al menos el coeficiente de correlación puede servir para confirmar el grado de la influencia “causal” de A en B . A continuación trataremos el tema desde el punto de vista de la “estadística”, tomando como base la definición de causalidad usada en el estudio de los modelos causales-estructurales (Bollen, 1989). Dicho autor muestra un resumen de las distintas ideas sobre la causalidad de distintos autores, como Blalock, Campbell y Stanley y las analiza para su posible empleo en la construcción de modelos de ecuaciones estructurales en ciencias sociales.

Un modelo de este tipo se basa en una serie de observaciones sobre el fenómeno estudiado, y delinea la asociación entre los factores o variables del fenómeno que interesan. Esta asociación, en general no es determinista, lo que no invalida capacidad pronóstica del modelo.

Supuestos

Supongamos una variable dependiente y_I aislada de toda posible influencia, salvo de la de otra variable independiente x_I . Si un cambio en x_I puede producir otro cambio (en algunos casos) en y_I , entonces, se dice que x_I es *causa débil* de y_I . En esta definición aparecen tres componentes, que son: aislamiento, asociación y dirección de la influencia (Batanero, 2001). Un problema es que el componente de aislamiento resulta casi inalcanzable en la realidad. Davis (1985) presenta tres características de una definición de *causación débil*, que son:

- Hablamos de tendencias, lo que hace que aparezcan excepciones. Por ejemplo, si se estudia mayor número de horas, generalmente (pero no siempre) se consigue mejor puntuación en un examen.
- Se dice “ x_I es una causa de y_I ”, ya que puede presentarse más causas para y_I . En el ejemplo, además del número de horas de estudio, el resultado del examen dependerá de la suerte con las preguntas propuestas, el nerviosismo del estudiante, etc.
- Únicamente la correlación no implica causación, como veremos posteriormente.

A continuación analizamos las ecuaciones que describen matemáticamente la

causalidad, desde el punto de vista determinista y aleatorio, siguiendo a Batanero (2001).

Modelo determinista lineal

La relación de causalidad se podría representar de forma matemática, mediante una dependencia funcional, que para una única variable dependiente y otra independiente se expresaría en la forma siguiente:

$$y_I = f(x_I)$$

Y en el caso de que la ecuación anterior sea lineal, con ordenada en el origen igual a cero, resultaría lo siguiente

$$y_I = \gamma_{II}x_I$$

Resultando que por cada unidad de cambio en x_I , se produciría exactamente γ_{II} unidades de cambio en y_I . En este caso aparece una sólo causa y un solo efecto (noción determinista de causalidad). Además se manifiesta la hipótesis de Hume de “conjunción constante”, es decir, por cada vez que ocurre la causa se produce el efecto (Estepa, 1993).

Modelo aleatorio lineal

Un modelo más realista según Bollen (1989), que introduciría el componente estocástico, sería el siguiente, que serviría para representar, por ejemplo, la relación entre el peso y_I , y la altura x_I de una persona; siendo γ_{II} el coeficiente de la ecuación de regresión.

$$y_I = \gamma_{II}x_I + \xi_I$$

En este caso ξ_I es un error aleatorio sin control directo, por lo que para cada caso particular no podemos evaluar su valor; generalmente se supone que tiene una distribución normal.

Aislamiento

Una condición según Davis (1985) para evaluar la relación de causalidad es que ni la variable dependiente ni la independiente estén afectadas por otras variables no controladas. Cuando nos planteamos este tipo de modelo haremos la hipótesis de independencia del término de error con el resto de las variables, permitiéndonos evaluar la influencia de x_I sobre y_I aisladamente de ξ_I , o sea, estimar el coeficiente de regresión.

La hipótesis de aislamiento se cumple cuando los datos se han tomado aleatoriamente, y han sido medidos todos con el mismo instrumento, pero puede no cumplirse en otros casos, por ejemplo en muestras no aleatorias (Batanero, 2001).

Modelos mas completos

Es verosímil que en un fenómeno real haya muchas causas de y_1 , además de x_1 . Un modelo del tipo que mostramos a continuación es mas realista, lo que supone que la correlación de cada variable explicativa con la independiente no será excesivamente alta y que más que de una sola causa habría que hablar de multiplicidad de causas.

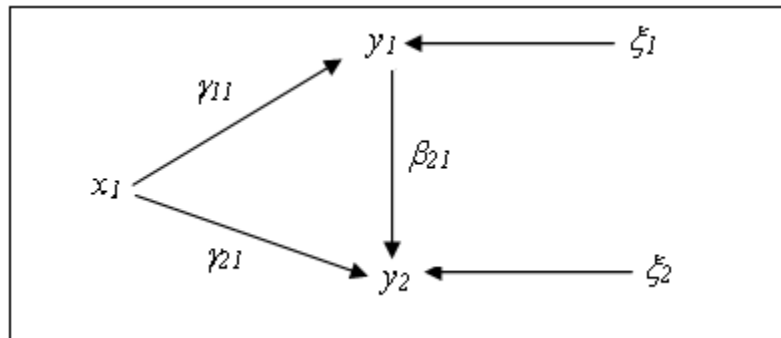
$$y_1 = \gamma_{11}x_1 + \gamma_{12}x_2 + \dots + \gamma_{1q}x_q + \xi_1$$

En este caso, el término de error ξ_1 es incorrelacionado con las variables x_i , $i=1, \dots, q$. Vamos a considerar el caso de incumplimiento de esta hipótesis, con dos ecuaciones:

$$y_1 = \gamma_{11}x_1 + \xi_1 \quad y_2 = \gamma_{21}y_1 + \gamma_{21}x_1 + \xi_2$$

estando ξ_1 y ξ_2 incorrelados entre sí y con x_1 (Figura 2.2.4.1, reproducida de Batanero, 2001, p. 34).

Figura 2.2.4.1. Diagrama de la situación descrita



En los casos erróneos en los que se omite la variable y_1 en la segunda ecuación, quedaría:

$$y_2 = \gamma_{21}x_1 + \xi_2^* \quad \text{con} \quad \xi_2^* = \beta_{21}y_1 + \xi_2$$

Está demostrado (Davis, 1985) que al estimar el efecto directo de x_1 sobre y_2 , o sea γ_{21} , obtendremos en lugar de esto, una estimación del efecto total $\gamma_{21}^* = \gamma_{21} + \beta_{21}\gamma_{11}$. Esto puede llevar a situaciones donde la correlación estimada entre dos variables sea mayor o menor que la existente en la realidad, o bien incluso se estime con un signo

contrario al de la correlacion verdadera, según muestra Estepa (1993).

Medidas de asociación propuestas por algunos filósofos sobre la causalidad

Algunos de los filosofos no solo han estudiado teoricamente las diferencias entre asociacion y causalidad, sino que han propuestos modos de medir la asociacion. En este apartado veremos cuatro medidas de asociación propuestas por distintos filósofos, que aparecen en el trabajo de Ellet y Erickson (1986). Suponemos que las variables son dicotómicas para poder hablar de signo de la asociación.

Kendall, Lazarsfeld y Nagel indican que dos variables A y B tienen asociación positiva si y solamente si la probabilidad de que ocurran simultáneamente A y B es mayor que el producto de las probabilidades de A y B :

$$P(A \cap B) - P(A)P(B) > 0$$

Los autores definen en forma parecida la asociación negativa y nula. En una tabla de contingencia esta definicion es equivalente a la diferencia entre frecuencias observadas y esperadas, que, si es positiva, daría lugar a una correlacion positiva, si es negativa a una negativa y si es nula a la independencia (Batanero, 2001).

Reinchebach y Suppes suponen que dos variables están asociadas positivamente si y solamente si la probabilidad de B condicionada a A menos la probabilidad de B es mayor que cero. Esto es, A y B están correlacionadas positivamente si y solamente si :

$$P(B/A) - P(B) > 0$$

Del mismo modo definen el caso de asociación negativa y nula. Estepa (1993) indica que esta definicion es equivalente a comparar la probabilidad condicional del suceso con la probabilidad simple. Sólo si son iguales las variables serían independientes.

Salmon y Suppes proponen que dos variables están asociadas positivamente si y solamente si la probabilidad de B condicionada a A menos la probabilidad de B condicionado a no A es mayor que cero. Esto es, A y B están correlacionadas positivamente si y solamente si:

$$P(B/A) - P(B/\bar{A}) > 0$$

Del mismo modo definen el caso de asociación negativa y nula. Estepa (1993) indica que esta definicion es equivalente a comparar la probabilidad condicional del suceso B condicionado con A con la probabilidad de B condicionado al contrario de A .

Finalmente, Asher y Blalock indican que dos variables están asociadas positivamente si y sólomente si el coeficiente Phi de Pearson es positivo; A y B están

asociadas negativamente si y sólo si el coeficiente Phi es negativo; y *A* y *B* no están asociadas si y solamente si el coeficiente Phi es cero.

Estepa (1993) indica que al desarrollar las formulas correspondientes a estas cuatro medidas el denominador no varia por lo que cualitativamente (en relacion al signo) son equivalentes. Desde un punto de vista cuantitativo (estimar la intensidad de la relacion) las medidas son iguales solo en caso de independencia de las variables pero en caso contrario pueden tomar diferente valor.

Sintesis

Como hemos visto las relaciones entre asociación y causalidad son complejas. Díaz y de la Fuente (2005b) analizan las relaciones entre causalidad y condicionamiento, que también aparecen en un problema de asociación. Indican que las personas construimos nuestro conocimiento del mundo sobre la base de relaciones de causa y efecto entre diferentes sucesos. Desde el punto de vista de la probabilidad, si un suceso *A* es la causa estricta de un suceso *B*, siempre que suceda *A*, sucederá *B*, por lo que $P(B/A)=1$ y habrá una asociación entre *A* y *B*. Pero también se puede establecer una condicionalidad en situaciones no causales, por ejemplo, en situaciones diagnósticas. En consecuencia, causalidad, correlación y condicionamiento están relacionados pero no son equivalentes, como se verá a lo largo de la memoria.

Sin embargo, Moreno (1998) considera que el esquema causal predominante heredado es insuficiente para diferenciar estas relaciones. Para concluir que dos o más factores tienen relación causa-efecto es necesario demostrar que la asociación entre estos es válida y causal, considerando una asociación válida cuando es una asociación real o verdadera, trayendo un efecto mínimo del azar y del sesgo, que está relacionado con el rigor metodológico del estudio. A pesar de un buen control interno, los resultados pueden no ser extrapolados o generalizados externamente. Esto ocurre cuando los grupos investigados no son representativos. Respecto a los diseños de investigación, se pueden clasificar, según su rigor metodológico decreciente en relación con la determinación de causalidad en: Ensayo controlado o estudio experimental; Estudio de cohortes; Estudio de casos y controles; Diseños transversales y Estudios descriptivos.

Para comprobar que la asociación es causal se deben considerar los siguientes aspectos (Álvarez y Pérez, 2004): fuerza de asociación, congruencia (si la asociación ha sido reproducida por investigadores diferentes), relación temporal, gradiente dosis-respuesta (a más magnitud más intensidad del efecto), credibilidad (concordancia de la

asociación con el conocimiento que se tiene en la actualidad), especificidad (genera un solo efecto).

Una vez clarificadas las relaciones entre asociación, condicionamiento y causalidad, pasamos a describir las investigaciones relacionadas con la forma en que las personas perciben la asociación, como paso previo al establecimiento de una relación causal entre las variables.

2.3. JUICIOS DE ASOCIACIÓN

A partir de una tabla de contingencia se pueden proponer gran variedad de problemas a los sujetos (Batanero y Díaz, 2008). Uno de estos problemas consiste en realizar juicios de asociación, es decir, decidir si las variables presentadas en filas y columnas tienen dependencia estadística o no la tienen.

Para McKenzie y Mikkelsen (2007) es difícil imaginar una actividad cognitiva más importante que la evaluación de la asociación. Algunas variables que se pueden considerar en un juicio de asociación son el tamaño de la muestra, que la distribución marginal de la variable independiente tenga o no la misma frecuencia para sus diferentes valores, la intensidad de la asociación, la dirección de la asociación, en caso de tabla 2x2 u ordinal, que concuerden la asociación empírica en la tabla con las creencias previas y la posible explicación de la asociación.

Los juicios de asociación han sido poco explorados en didáctica de las matemáticas, exceptuando los trabajos realizados por Estepa y sus colaboradores, que se describen más adelante. Por el contrario, el problema resulta de gran interés en psicología, y está relacionado con la toma de decisiones en ambiente de incertidumbre (Scholz, 1987) porque en muchas de estas decisiones se necesita un juicio sobre la asociación de las variables. La mayor parte de estos estudios sólo investigan las tablas 2x2 (dos filas y dos columnas). En lo que sigue se analizan estos trabajos, comenzando por el pionero de Inhelder y Piaget.

2.3.1. DESARROLLO EVOLUTIVO DEL CONCEPTO DE ASOCIACIÓN

Los que iniciaron el avance de la investigación sobre la asociación fueron Inhelder y Piaget (1955), que consideraron la comprensión de la asociación como el último paso en el desarrollo de la comprensión de la probabilidad, como exponen en el capítulo XV de su obra “De la lógica del niño a la lógica del adolescente”. Así que los antecedentes evolutivos de los conceptos de asociación y probabilidad son iguales para

estos autores, quienes sugieren que para entender lo que es la asociación es necesario comprender la proporción, combinatoria y probabilidad. Para exponer mejor sus estudios, comenzamos realizando un resumen de la teoría de Piaget sobre etapas en el desarrollo evolutivo de la inteligencia.

Etapas del desarrollo evolutivo según Jean Piaget

Desde que nacemos hasta que llegamos a la vida adulta, nuestra forma de pensar, así como el razonamiento matemático van evolucionando. La teoría desarrollada por Piaget (1975) indica que cuando un individuo afronta un problema matemático, lo intenta resolver mediante los conocimientos que ya posee, usando esquemas conceptuales existentes. Como resultado de la asimilación, el esquema cognitivo existente se reconstruye o expande para acomodar la situación. Un niño va comprendiendo su entorno del siguiente modo: (a) mejorando su sensibilidad a las contradicciones; (b) realizando operaciones mentales; (c) comprendiendo las transformaciones; y (d) adquiriendo nuevas nociones.

Las etapas, según Piaget (1975), que determinan el desarrollo evolutivo son las siguientes (ver, por ejemplo, Piaget e Inhelder, 1951, para la descripción de estas etapas en el campo de la probabilidad):

- *Período sensoriomotor* (0-2 años). Comienza la manipulación de objetos; se perciben y experimentan propiedades (color, tamaño, forma, textura, sabor, olor...); a los 5 meses se discrimina conjuntos de 2-3 ítems; a los 10 meses se discrimina conjuntos de 3-4 ítems.
- *Período preoperacional* (2-7 años). Organiza el espacio, situando y desplazando los objetos (dentro/fuera, encima/debajo, delante/detrás, arriba/abajo); descubre y compara propiedades físicas de los objetos que manipula: longitud, distancia, cantidad; utiliza diferentes formas de etiquetado para diferenciar colecciones numéricas de pocos elementos; contrasta magnitudes por comparación y estima, a partir de una cantidad, la longitud, volumen y peso; ordena en el tiempo; engloba aspectos de tipo espacial; se inicia en el conteo; trabaja con una sola cantidad y resuelve problemas de cambio sencillo.
- *Período de las operaciones concretas* (7-11). Aparición de operaciones reversibles con la adquisición de principios de conservación de cantidad, peso y volumen; compara y cuantifica mediante la geometría y el sistema métrico decimal y

representa datos gráficamente; agrupa los objetos en función de propiedades aditivas o multiplicativas; ordena elementos en función de la cualidad que varía; adquiere la noción de sistema de numeración y de operación con números; comprende conceptos espaciales: espacio que ocupan los objetos y su desplazamiento (topológicas, proyectivas euclidianas, métricas...); y, operaciones temporales y cinéticas: orden de sucesión de los objetos en el espacio.

- *Período de operaciones formales* (11-15). El período de las operaciones formales constituye el último del desarrollo intelectual del hombre, atribuyéndosele la máxima importancia en el desarrollo de los procesos cognitivos y sociales (Inhelder y Piaget, 1955; Piaget, 1975). Características del pensamiento formal son: (a) se contempla lo real como parte de lo posible, (b) se acentúa lo hipotético-deductivo frente a lo, empírico-inductivo (c) se depura el pensamiento proposicional, (d) se acentúa la diferencia entre inteligencia práctica y especulativa, (e) se incrementa la cantidad y calidad de las estrategias de procesamiento de la información, (f) se potencia y acentúa el análisis crítico frente a las percepciones globales, (g) se depura y da carácter sistemático al método de análisis, (h) se desarrollan y amplían las posibilidades combinatorias de pensamiento, (i) se da la integración de las dos formas de pensamiento reversible: la inversión y la compensación o reciprocidad.

Según Inhelder y Piaget (1955), la adquisición de las operaciones formales viene representada estructuralmente por el uso del constructo conocido por el nombre de "grupo INRC", siendo: I (identidad), N (negación o inversión), R (reciprocidad), C (correlación, es decir lo que nosotros denominamos asociación). Este grupo se caracteriza por las siguientes propiedades: composición (combinar elementos del mismo conjunto con otro elemento que también pertenece al conjunto); asociatividad (la combinación de elementos de un conjunto no depende de la forma de agruparse); identidad general (existe un sólo elemento que al combinarse con otro lo hace idéntico, elemento identidad); y reversibilidad (por cada elemento de un conjunto existe otro que combinado con él resulta el elemento identidad).

Además, las operaciones específicas de este grupo son: identidad (consistente en no variar una determinada proposición); negación (consistente en la inversión de la proposición idéntica); reciprocidad (consistente en la producción del mismo efecto que la operación idéntica pero en otro sistema); y correlatividad (consistente en la negación o inversión de la operación anterior).

Un último aspecto, en cuanto a estructura en este periodo, es el de los ocho esquemas de operaciones formales, que son (Piaget, 1975): (a) esquema combinatorio (combinaciones, variaciones, permutaciones), (b) esquema de proporcionalidad (asimilación de que $X/Y=X'/Y'$ con tal que $X.Y'=X'.Y$), (c) esquema de doble referencia (consistente en la coordinación de dos sistemas de referencia y la relatividad de los movimientos o las velocidades), (d) esquema de equilibrio mecánico (toda acción le corresponde una reacción de la misma intensidad pero en sentido contrario), (e) esquema de probabilidad (puesta en marcha de la predicción de eventos), (f) esquema de correlación (comprensión de la variación conjunta de dos o más variables), (g) esquema de compensaciones multiplicativas (asimilación de la conservación de volumen), y (h) esquema de conservación que va más allá de la materia (aplicación de la conservación a partir de razonamientos deductivos o inductivos, como movimiento rectilíneo y uniforme, etc.).

En lo que sigue analizamos con mayor detalle el esquema de asociación (que Inhelder y Piaget denominan correlación).

El esquema de la correlación (asociación)

Para estudiar la comprensión de la asociación, Inhelder y Piaget (1955) realizaron un estudio con chicos a partir de 12-13 años, donde clasificaron las posibilidades de combinar el color de pelo (rubio o moreno) y color de ojos (azules o negros), usando como material unas cuarenta tarjetas con dibujos coloreados, donde aparecían caras en las que se veían el color de ojos y color de pelo. Se podía hacer varias pruebas con los sujetos, (a) darle al sujeto unas pocas tarjetas sin clasificar para que dijera si existe alguna relación entre las variables, (b) darle unas pocas tarjetas clasificadas para deducir si existe alguna relación entre las variables, o (c) darle dos grupos de tarjetas para que dijera en qué grupo la relación era más nítida.

Tabla 2.3.1. Esquema de una tabla de contingencia 2x2

	B	\bar{B}	Total
A	a	b	$a+b$
\bar{A}	c	d	$c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

Para describir sus resultados mostraron sus datos con un esquema de la tabla de contingencia (Tabla 2.3.1). En este esquema, los valores interiores, es decir los valores

a , b , c y d , son las frecuencias absolutas de cabello rubio y ojos azules, cabello rubio y ojos negros, cabello moreno y ojos azules, y cabello moreno y ojos negros.

En su estudio los autores indican que, para estudiar una asociación entre las variables, había que tener en cuenta los casos favorables, correspondientes a $(a+d)$, los casos desfavorables, correspondientes a $(b+c)$ y los casos posibles, $(a+d)+(b+c)$. Para los autores, la asociación aparece cuando el número de casos favorables es mayor que el de los desfavorables.

Según Piaget e Inhelder, los chicos de entre 11-12 años, estiman probabilidades simples, pero solamente comparan celdas dos a dos, entendiendo la presencia de relación entre las variables, pero sin entender que a y d tienen la misma significación. Los adolescentes de su estudio que están en el primer nivel de comprensión de la asociación (que corresponde al nivel *IIIa* en el desarrollo del concepto de probabilidad), sólo se fijaban en los casos positivo-positivo, o sea la relación del caso a al total de datos, sin mirar el caso negativo-negativo, es decir el caso d (Tabla 2.3.2.).

Por tanto, sólo usan una de las cuatro variables de la tarea. Un segundo nivel de razonamiento (nivel *IIIb* en el desarrollo del concepto de probabilidad) en relación a la asociación es cuando los adolescentes utilizan dos de las variables y comparaban a con b , ó a con c , nuevamente sin mirar el caso negativo-negativo, es decir el caso d , sin darse cuenta de que los casos de la celda d tienen el mismo valor normativo que los de la celda a .

Además, al comprender cuales son los casos favorables (a y d) y desfavorables (b y c) de la asociación, no los comparan. Según los autores el problema reside en comprender que la asociación no se puede estimar calculando una simple probabilidad, es decir, una relación elemental entre los casos favorables $(a+d)$ y el conjunto de los casos posibles $(a+d+b+c)$. Ello explica por qué la comprensión de la asociación sólo se adquiere en el nivel *IIIb*, aunque se posea la comprensión probabilidad simple en el nivel *IIIa* (Inhelder y Piaget, 1955).

Al final de esta etapa (14-15 años) se comienza a establecer las relaciones diagonales, comparándolas entre si o con el total $(a+b+c+d)$. Los sujetos los relacionan directamente $(a+d)$ con $(b+c)$ para evaluar cuanta correlación hay, viendo los distintos tipos de correlación: positiva $(a+d) > (b+c)$, nula $(a+d) = (b+c)$ y negativa $(a+d) < (b+c)$.

Tabla 2.3.2. Significado de las celdas en una tabla 2x2

	B	\bar{B}	Total
A	Positivo-positivo	Positivo-negativo	Suma A
\bar{A}	Negativo-positivo	Negativo-negativo	Suma \bar{A}
Total	Suma B	Suma \bar{B}	Suma A, B, \bar{A} y \bar{B}

Inhelder y Piaget pensaron que para llegar a la comprensión completa de la asociación se debía utilizar la siguiente fórmula que proporciona una evaluación de la asociación entre distintas variables. También indica que no se suele usar esta relación hasta cerca de los 15 años de edad en el estadio *IIIb*.

$$R = \frac{(a + d) - (b + c)}{(a + d) + (b + c)}$$

Si las probabilidades $\frac{(a + d)}{(a + d) + (b + c)}$ y $\frac{(b + c)}{(a + d) + (b + c)}$ son iguales, entonces

$R=0$, lo que implica que la asociación sería cero, siendo más fuerte cuanto mayor fuese a diferencia $(a+d)-(b+c)$. Esto quiere decir que los autores sólo consideraron la asociación positiva en su trabajo, aunque al final de su estudio comentan algo sobre correlación negativa. Shaklee y Tucker (1980) informan que esta estrategia, al intentar identificar el nivel de intensidad de la asociación, puede dar lugar a errores:

“Esta estrategia identificará correctamente la mayoría de las contingencias, pero puede producir errores, cuando las frecuencias de los estados alternativos del Factor A (y/o el Factor B) son muy diferentes” (p. 460).

2.3.2. ESTRATEGIAS EN LOS JUICIOS DE ASOCIACIÓN

Después de Inhelder y Piaget fueron muchos los estudios sobre la percepción de la asociación, aunque la mayoría se realizaron en el campo de la psicología. El trabajo de Estepa (1993) y otros posteriores de este autor abren el tema dentro de la Didáctica de la Matemática. En lo que sigue se resumen estas investigaciones.

2.3.2.1. INVESTIGACIONES EN PSICOLOGÍA

Autores como Smedlund (1963) y Shaklee y Mins (1982) observaron sujetos adultos que, para realizar un juicio de asociación, sólo se fijaban en la celda primera de la tabla de contingencia (presencia de los dos caracteres) o que la comparaban con la segunda celda (presencia de uno y ausencia del otro). Es decir, a pesar de hallarse en el periodo de las operaciones formales, siguen razonando como los niños de etapas anteriores. En consecuencia, estos sujetos no utilizaban todos los datos relevantes en el

problema, por lo que, en contra de las teorías de Inhelder y Piaget, el desarrollo del razonamiento adecuado sobre concepto de asociación no sería espontáneo en los sujetos, a menos que se dé una instrucción formal sobre el mismo.

Jenkins y Ward (1965) explican estos errores por la dificultad que conllevan las tareas relacionadas con emitir un juicio de asociación. Los autores observan que, hasta la estrategia de comparar las diagonales, que fue considerada correcta por Piaget e Inhelder, sólo se puede usar para emitir un juicio de asociación cuando las frecuencias marginales sean iguales para la variable independiente, pues en otros casos puede dar lugar a un juicio de asociación incorrecto. Aunque esto ocurra, esta estrategia es utilizada por adultos en todo tipos de tablas (incluso en aquellas en que lleva a un error en el juicio de asociación) en las investigaciones de Allan y Jenkins (1983) y Shaklee y Tucker (1980).

Jenkins y Ward proponen para casos generales como estrategia correcta comparar la diferencia entre las probabilidades de que se dé una variable en presencia o ausencia de la otra, es decir $P(B/A)$ y $P(B/\bar{A})$. Dicha comparación podría hacerse mediante la diferencia de estas dos probabilidades:

$$P = \left(\frac{a}{a+b}\right) - \left(\frac{c}{c+d}\right)$$

En la investigación psicológica posterior sobre juicios de asociación, se observan una variedad de estrategias, las cuales se pueden agrupar en niveles de complejidad. Pérez Echeverría (1990) describe los siguientes niveles, basándose en el número de celdas utilizadas y describe los porcentajes de uso por estudiantes de psicología en su estudio:

- Nivel 1. Utilizar sólo una celda, generalmente la celda a (uso 3,7%).
- Nivel 2. Comparar la celda a con b o c , es decir usar dos datos (uso 30,56%).
- Nivel 3. Comparar la celda a con b y c , es decir usar tres datos (uso 15,74%).
- Nivel 4. Usar todos los datos pero sólo mediante comparaciones aditivas (uso 37,96%).
- Nivel 5. Concluir la asociación por el uso de una relación multiplicativa entre las cuatro celdas (uso 12,04%).

En nuestro trabajo haremos un estudio más completo de las estrategias utilizadas, basándonos no solamente en el número de celdas, sino en los objetos matemáticos implícitos en cada estrategia y en su corrección. No obstante, en el estudio

de evaluación inicial (Capítulo 3) se analizará cada uno de los niveles de Pérez Echeverría (1990), así como los tipos de estrategias descritos en esta sección para estudiar la frecuencia de uso en los participantes.

2.3.2.2.SESGOS EN LA DETECCIÓN DE LA ASOCIACIÓN Y EFECTO DE LAS TEORÍAS PREVIAS

Otro problema diferente, relacionado con los juicios de asociación y que ha sido muy investigado en las últimas décadas es la valoración de las asociaciones selectivas relevantes en humanos. La idea principal que ha guiado esta línea de investigación es que estas asociaciones no tienen por qué estar restringidas al dominio del condicionamiento clásico, sino que pueden ser una manifestación de un esquema más general, que induce sesgos de covariación.

Para valorar los sesgos de covariación, se ha empleado el paradigma de *correlación ilusoria*. Este constructo, desarrollado por Chapman y Chapman (1967, 1969), implica la exposición a diferentes categorías de eventos que no están correlacionados. Después de la exposición a la serie de estímulos y resultados, se pide a los sujetos que realicen juicios respecto al grado de covariación establecido entre ellos.

Chapman y Chapman (1969) describen la presencia en muchas personas de este sesgo que consiste en formarse teorías sobre la relación entre variables que impide evaluar correctamente las contingencias empíricas. En algunas ocasiones el cerebro tiende a interpretar la realidad de manera sesgada, creyendo detectar relaciones que en realidad son inexistentes. Por ejemplo, si dos eventos van frecuentemente juntos, es probable que se concluya que existe algún tipo de relación entre ellos, incluso cuando esto sea falso. Se habla entonces de "*correlaciones ilusoria*", que se define como la percepción de una relación o dependencia entre elementos que en realidad son independientes. Pérez Echeverría (1990, p. 101) reproducen la siguiente definición de "*correlación ilusoria*" dada por Murphy y Medin (1985):

"Cuando percibimos que existe una correlación basándonos en nuestras propias teorías pero no existe ningún hecho empírico que las sustente, se habla de correlación ilusoria".

La experiencia vital y el entorno cultural son los responsables de estas teorías, porque dichas variables llaman la atención de las personas. Un ejemplo sería el encontrar una falsa asociación entre la pertenencia a un grupo minoritario y los comportamientos del grupo.

Chapman y Chapman estudiaron este efecto en pruebas psicodiagnósticas. Su estudio demostró que aunque la prueba descriptiva no sea útil en la diagnosis de los trastornos mentales, algunos psicólogos las utilizaban, debido a la correlación ilusoria, entre los resultados de la prueba y ciertos diagnósticos. Un ejemplo de una prueba descriptiva es "Dibujar una persona", donde se pide a los pacientes dibujen a persona en un trozo de papel en blanco. Algunos psicólogos creen en una correlación entre el dibujo de una persona con los ojos grandes y la paranoia, aunque tal correlación no existe.

Muchos otros autores han estudiado la influencia de las teorías previas en el contexto del problema en los juicios de asociación (Jennings, Amabile y Ross, 1982; Wrigth y Murphy, 1984; Alloy y Tabacnick, 1984). En términos generales se puede decir, que cuando los datos no reflejan los resultados esperados por estas teorías, aparece en los sujetos un conflicto cognitivo.

Algunos de estos sesgos han sido atribuidos a la *heurística de la disponibilidad*, que consiste en estimar la probabilidad de ocurrencia de determinados acontecimientos basándose en la mayor o menor facilidad con que puede construir ejemplos del mismo (Tversky y Kahneman, 1974). Esta heurística, generalmente asociada a experiencias personales, puede llevar a sesgos en el razonamiento probabilístico, incluido la estimación de la asociación. Además de la correlación ilusoria, otros mecanismos a que da origen esta heurística son los siguientes:

- *Recuerdo de ejemplos.* Un ejemplo que se ha presenciado personalmente – por ejemplo un accidente- se recuerda mejor y se juzga más probable.
- *Sesgo debido a la efectividad de patrones de búsqueda.* Si se piensa por un momento en una palabra en español de tres o más letras. ¿Es más probable que una palabra escogida al azar y que contenga una letra “r” comenzaría con la “r”, o la “r” sería la tercera letra?
- *Sesgos debido a la imaginación.* Cuando se juzga, por ejemplo, más peligroso un viaje en avión que un viaje en coche.

Tversky y Kahneman (1974) sostienen que la disponibilidad es producto de la percepción incorrecta de la frecuencia de un suceso, lo que puede influir en las decisiones que se tomen en ciertas circunstancias. Debido a ello, distintos tipos de profesionales podrían tomar decisiones importantes en su campo de trabajo, únicamente sobre la base de sus recuerdos más significativos.

Kahneman, Slovic y Tversky (1982) son poco optimistas sobre los efectos de la enseñanza sobre las heurísticas y los sesgos, ya que consideran que estos son muy difíciles de cambiar. Ellos sugieren que hay una interferencia psicológica que lleva a las personas, incluso con formación en probabilidad, a dar respuestas subjetivas, basadas en heurísticas y sesgos. Una posición distinta tiene Sedlemeier (1999) quien señala que la enseñanza mediante la resolución de problemas puede ayudar a mejorar la comprensión de tópicos de probabilidad y estadística.

2.3.2.3. INVESTIGACIÓN EN DIDÁCTICA DE LA MATEMÁTICA

Como se ha indicado, la investigación sobre las tablas de contingencia ha sido muy escasa dentro de la didáctica de la matemática. En España la investigación sobre tablas de contingencia, fue llevada a cabo especialmente por Antonio Estepa en diversos trabajos. En su tesis doctoral (Estepa, 1993) aborda un estudio histórico-epistemológico de la asociación estadística; un estudio sobre concepciones previas en universitarios sobre la asociación; una experiencia de enseñanza de la asociación, basado en la resolución de problemas y el uso de ordenador; el estudio de la evolución de las concepciones iniciales después de la enseñanza y de las dificultades encontradas. Posteriormente, Sánchez-Cobo realiza una investigación de libros de texto de Bachillerato sobre la presentación de los temas de correlación y regresión (Sánchez, 1996) y un estudio de evaluación sobre la correlación en primeros cursos universitarios. (Sánchez, 1999). Nosotros sólo tendremos en cuenta la parte de la investigación centrada en las tablas de contingencia.

Estepa (1993) y posteriormente Estepa y Batanero (1995) analizan los juicios de asociación, además de las estrategias empleadas por alumnos del Curso de Orientación Universitaria de tres institutos de Bachillerato, trabajando con tablas de contingencia 2×2 y tablas $r \times c$ (r filas y c columnas). Nos muestran las frecuencias de juicios correctos e incorrectos, diferencia de dificultad de algunos ítems y un análisis factorial de las respuestas en los juicios de asociación.

La muestra seleccionada de forma intencional, estaba formada por 213 alumnos que se distribuían de la siguiente forma: 124 de COU de ciencias y 89 de COU orientado a letras. Correspondiendo 113 a varones y 100 a mujeres. Treinta y cuatro alumnos tenían conocimientos previos de estadística y sólo tres los que sabían algo de dependencia funcional y aleatoria. El fin de estudio fue realizar una evaluación sobre los conocimientos iniciales de asociación estadística en alumnos que ingresan a la

universidad. Además, el cuestionario se comprobó en una muestra piloto de 51 estudiantes, obteniéndose un índice de generalizabilidad de 0,84.

El cuestionario utilizado consta de diez preguntas donde hay que dar un juicio sobre la asociación de dos variables: cinco de las preguntas mostrando los datos en forma de tabla de contingencia, tres preguntas usando diagrama de dispersión y dos preguntas de comparar una variable numérica en dos muestras. Las preguntas de tabla de contingencia iniciales, se dividían en 3 sobre tablas 2x2, una cuarta sobre una tabla 2x3 y una última sobre una tabla 3x3.

Estos problemas fueron extraídos de otras investigaciones, de datos reales o contruados para la investigación. Para la construcción se tuvieron en cuenta las variables de tarea siguientes: tipo de dependencia entre las variables (directa, inversa e independencia), intensidad de la dependencia (midiéndolo con el coeficiente de Pearson, el coeficiente PHI o el coeficiente de Cramer), diferenciación de la variable dependiente e independiente (en unos casos se ve claramente y en otros no), concordancia entre los datos y las teorías previas (en ocasiones los datos coinciden con las teorías previas y otras no), y tipo de covariación según Barbancho (1973) (dependencia causal unilateral, interdependencia, dependencia indirecta, concordancia y covariación casual).

En uno de los ítems sobre tablas de contingencia se observa cómo hay sujetos de la prueba que consideran la asociación inversa como independencia. También cabe destacar la mayor dificultad de los ítems en que los datos presentan una relación de las variables distinta a la que cabe esperar por las teorías previas.

Se espera relación entre las respuestas de un mismo sujeto, así que para determinar la estructura, realizan un análisis factorial (utilizando BMDP), con el método factor principal y como método de rotación la rotación varimax. A través de la estructura de los autovalores se confirma la multidimensionalidad del juicio de asociación. Los factores con coeficiente factoriales de magnitud apreciable se interpretan en la forma siguiente: asociación entre variables como variación de las frecuencias condicionales (agrupando las tablas de contingencia, salvo el ítem que presenta asociación inversa), asociación entre variables como variación del valor numérico de una de ellas en función de la otra variable (agrupa el ítem 7), signo de la asociación cuando no hay teorías en contra (opone el ítem 3 frente los ítems 5, 6 y 7), concordancia frente a causalidad (correlación moderada), y asociación como comparación de muestras independientes (agrupa con signos opuestos las tablas 2x2 y el último ítem de comparación de muestras independientes). Estos factores con inercia

superior a la unidad, explican un 62 % de la varianza, teniendo todos fuerte especificidad.

Los autores muestran un estudio cualitativo de estrategias de los 213 estudiantes, analizando su corrección, desde el punto de vista de los conceptos matemáticos que se usan. Para el análisis de estrategias, Estepa y Batanero aplican el marco conceptual de Vergnaud (1990) quien describe dos tipos de situaciones: situaciones respecto a las cuales los sujetos tienen las competencias necesarias, y situaciones en las que el sujeto no tiene estas competencias. En el estudio de Estepa y Batanero, los sujetos, estarían en la segunda situación descrita. Vergnaud utiliza la palabra esquema para definir el proceso de organización de la conducta que se emplea en cada una de estas dos situaciones, observándose en la primera situación un esquema único y bastante automático, mientras que en la segunda situación se utilizaran varios esquemas hasta llegar a la solución del problema. Vergnaud describe los “*teoremas en acto*” que no son propiamente teoremas que aparecen en los libros de texto, sino propiedades y relaciones de los objetos matemáticos usados implícitamente al resolver los problemas.

En la investigación de Estepa se identifican estrategias empleadas en tablas de contingencia en diagramas de dispersión y en la comparación de muestras, y, dentro de cada tipo, estrategias correctas, parcialmente correctas (que proporciona soluciones válidas en algunos problemas, pero en otros no) e incorrectas. Describimos a continuación las usadas en las tablas de contingencia, junto con los teoremas en acto utilizados.

Estrategias correctas

- ST.1. *Comparar todas las distribuciones de frecuencias relativas condicionales de una variable para los distintos valores de la otra variable.* El alumno emplearía implícitamente el siguiente teorema en acto T1: La dependencia de una variable B respecto a otra variable A lleva a la variación de las frecuencias relativas condicionales $h(B/A)$ cuando A cambia. Una interpretación similar tendría el caso en que el alumno intercambiase filas por columnas.
- ST.2. *Comparar todas las frecuencias relativas condicionadas de una variable para un único valor con la marginal correspondiente de la otra variable.* En este caso el alumno emplearía implícitamente el siguiente teorema en acto T2: La distribución de B al condicionar por A , siendo independientes, implica la invarianza de la

distribución de B .

- ST.3. *Comparación de posibilidades, comparando las frecuencias de casos a favor y en contra de B en cada valor de A .* En esta estrategia se emplearía implícitamente el siguiente teorema en acto T3: Existe una correspondencia unívoca entre la probabilidad de un suceso y la razón de sus probabilidades a favor y en contra.

Estrategias parcialmente correctas

- ST.4. *Comparar la distribución de frecuencias absolutas condicionales con la frecuencia absoluta marginal correspondiente.* Esta estrategia supone utilizar comparaciones aditivas o cualitativas, de tal forma que no se cuantifican bien las probabilidades usadas, aunque es parecido a la estrategia ST.1. La estrategia no es correcta pues se comparan frecuencias absolutas y no relativas.
- ST.5. *Comparar sólo una frecuencia absoluta en cada distribución condicional, con la frecuencia marginal que le corresponde.* En este caso el alumno emplearía implícitamente el siguiente teorema en acto T4: con la variación de la frecuencia relativa de algún valor de A cuando B cambia se demuestra la dependencia. Tampoco en este caso la estrategia es correcta, pues se comparan frecuencias absolutas y no relativas.
- ST.6. *Comparar las frecuencias absolutas condicionadas la una con la otra.* Esta estrategia es parecida a la ST.2, aunque se comparan las frecuencias absolutas entre sí, por lo que no es correcta.
- ST.7. *Comparar la suma de frecuencias en las diagonales.* En este caso el alumno usa la estrategia descrita por Piaget e Inhelder, siendo correcta, sólo en el caso de que la tabla tenga igual las frecuencias marginales para la variable independiente.

Estrategias incorrectas

- ST.8. *El uso único de la celda de mayor frecuencia.* El alumno no utiliza toda la información presente en el problema, tan sólo la celda más sobresaliente.
- ST.9. *El uso de sólo una distribución condicional.* En esta estrategia el alumno no ve el problema como un problema de comparación de probabilidades.
- ST.10. *Comparar frecuencias dobles con el número total de observaciones.* En este caso el alumno, de forma particular elige las celdas de mayor frecuencia.

- ST.11. *Comparar frecuencias marginales*. Esta estrategia es incorrecta, aunque el alumno cree que obtener distintas frecuencias marginales, implica no poder llegar a una conclusión.
- ST.12. *Otros procedimientos incorrectos*. En casos esporádicos se usan procedimientos, o no relacionados con las frecuencias de la tabla, por ejemplo plantear una ecuación.

2.3.3. CONCEPCIONES SOBRE LA ASOCIACIÓN

Estepa (1993) cree que las estrategias usadas por los estudiantes en los juicios de asociación dependen de algunas concepciones que poseen los alumnos sobre la misma, describiendo la concepción causal, determinista, unidireccional y local, las cuales comentaremos a continuación.

- *Concepción causal*: Cuando el sujeto sólo considera la dependencia entre variables si puede adjudicarse a la presencia de una relación causal entre las mismas.
- *Concepción determinista*: Este término describe el caso de que los sujetos no admiten excepciones en la relación, implicando esto que a cada valor de la variable independiente le corresponde un solo valor de la variable dependiente. Este caso se presenta cuando los sujetos afirman que no hay asociación por el hecho de que en una tabla 2x2 hay casos en las celdas *b* y/o *c*. Otro claro ejemplo es cuando el sujeto exige la existencia de una expresión algebraica que relacione las variables.
- *Concepción unidireccional*: En este caso el estudiante no admite la asociación inversa, considerándose la intensidad de la asociación, pero no su signo. Presentándose casos en los que se dice que la dependencia inversa es independencia.
- *Concepción local*: Esta concepción se presenta cuando los sujetos, dan su solución mirando únicamente algunos casos aislados. Siendo ejemplos cuando sólo se tienen en cuenta los casos que confirman la asociación, observando una sola distribución condicional o fijándose en la celda de máxima frecuencia.

2.3.4. EVOLUCIÓN DESPUÉS DE LA ENSEÑANZA

Estepa (1993) realiza también un estudio del efecto de la enseñanza sobre el aprendizaje del concepto de asociación. La muestra estuvo formada por 22 alumnos de tercer curso de la Escuela Universitaria de Formación del Profesorado de E.G.B. de Jaén, de los cuales 17 no habían estudiado Estadística con anterioridad, 4 si lo habían

estudiado y uno no cumplimentó el pretest. Se puede considerar que se está en la segunda situación propuesta de Vergnaud (1990).

En la evaluación realizada al finalizar la enseñanza, el cuestionario estaba formado con 3 ítems de tablas de contingencia (tablas 2x2, 2x3 y 3x3), 3 ítems sobre la nube de puntos y 2 ítems sobre comparación de muestras (una dicotómica y otra numérica). Se realizó un estudio paralelo al de la muestra de comparación descrita en el apartado anterior para poder situar a los alumnos en un contexto más amplio.

En los ítems de tablas de contingencia, los alumnos leían las tablas sin dificultad, surgiendo los problemas al intentar calcular frecuencias relativas condicionadas y marginales. Respecto a los juicios de asociación, se puede observar cómo detectan los alumnos el tipo de asociación existente en una tabla 2x2. Algunos alumnos fallan por falta de razonamiento proporcional o por imponer sus teorías previas. En un ítem de dependencia inversa se observa la concepción unidireccional de la asociación en este tipo de tablas.

2.3.4.1. EVOLUCIÓN DE ESTRATEGIAS EN JUICIOS DE ASOCIACIÓN

Para el estudio de las estrategias, Estepa utiliza el método del autoinforme, como Smedlund (1963) o Beyth-Marom (1982), pues los alumnos responden de forma escrita, y el autor hace un análisis de contenido de dichos informes. Se clasificaron las respuestas en procedimientos o maneras diferentes de resolver el problema, observando la información empleada en la respuesta y al modo en que se compara o combina, y se realiza un recuento de frecuencias. A continuación agrupa cada procedimiento en grupos de estrategia similar.

La primera clasificación utilizada fue considerando estrategias diferentes según las celdas que se habían usado en la resolución del problema y el modo de utilizar la información. También presenta un apartado de “otras”, que incluye los procedimientos que no se pueden incluir en los apartados anteriores. Concluye lo siguiente:

- Los alumnos tienden a usar 1, 2 o 4 celdas, y nunca emplean 3 celdas.
- Un alto porcentaje de alumnos usan 1 celda, generalmente la de máxima frecuencia.
- Aproximadamente un tercio usan 2 celdas de una condicional, comparando con la marginal o entre sí.
- Los alumnos usan procedimientos aditivos, más que multiplicativos, para realizar un juicio de asociación correcto en tablas 2x2.

En Batanero, Estepa y Godino (1997) realizan un estudio de los cambios después de la instrucción, mediante un cuestionario, observándose como había influido la instrucción en sus falsas ideas y estrategias incorrectas. En las tablas de contingencia se aprecia que el número de estrategias correctas cambió de 17 a 31. Sin embargo, algunos alumnos que anteriormente a la instrucción utilizaron estrategias correctas, después emplearon estrategias incorrectas.. De forma general se observa una mejora, ya que la diferencia de puntuaciones total es positiva. Además se ve la variabilidad en las puntuaciones de los estudiantes y en los diferentes problemas.

Un ítem que no presentaba mejora, fue el de asociación inversa en una tabla 2x2, pero los alumnos identificaron la independencia y extendieron el uso de las estrategias correctas en tablas de contingencia 2x2 a tablas rxc.

2.3.4.2. JUICIOS DE ASOCIACIÓN USANDO ORDENADOR

En su investigación se usan ordenadores como recurso didáctico (Batanero, Estepa y Godino, 1991). Este uso, permite en el caso de la estadística el poder procesar datos, poniendo a los alumnos ante situaciones-problemas relativos al análisis de ficheros de datos de aplicaciones reales. Estos autores utilizan la técnica del registro de la interacción alumno-ordenador, viéndose los procedimientos empleados en la resolución de las cuestiones que se les plantean. También usando argumentación escrita en que basan sus respuestas, para ver la corrección y pertinencia de los procedimientos empleados.

Los alumnos durante la enseñanza trabajaron con un software didáctico preparado para la experiencia denominado PRODEST. Obteniéndose los ficheros de datos por los alumnos utilizando una encuesta a sus compañeros, o bien fueron suministrados por el profesor. Los programas usados fueron: CONTAJE (tablas y gráficos de datos categóricos), HISTO (tablas y gráficos de datos agrupados), TRONCO (gráfico del tallo y hojas), MEDIANA (cálculo de la mediana), ESTADIS (cálculo de estadísticos), TABLAS (análisis de tablas de contingencia) y REGRESIO (regresión y correlación bivariante).

La enseñanza alternó clases tradicionales en el aula ordinaria y clases prácticas en la sala de informática. En cada práctica, se presenta una situación problemática de análisis de un conjunto de datos, donde se muestran una serie de preguntas, que podrían

solucionarse con el uso de alguno de los programas mencionado en el párrafo anterior.

Clasificándose los problemas en los tipos siguientes:

- Análisis de una sola variable en el fichero completo.
- Análisis de una variable en una parte del fichero.
- Estudio conjunto de dos variables.
- Comparación de muestras independientes.
- Comparación de muestras relacionadas.
- Seleccionar la solución más adecuada entre varias opciones.
- Comprobación de alguna propiedad matemática de distribuciones y sus estadísticos.
- Plantear preguntas en un conjunto de datos que se solucionen con el uso de un procedimiento estadístico.

Se observa el refuerzo del estudio de la asociación entre variables, respecto a la enseñanza tradicional. También hay que mencionar lo esencial que es el tipo de variable estadística (dicotómica, cualitativa o cuantitativa), además de ciertas características de su distribución (necesidad de agrupación, ser o no simétrica, tener valores atípicos en caso de asimetría y bimodalidad).

Con el uso de grabaciones de la actuación del alumno con el ordenador se observó el aprendizaje, y una evaluación final individual de cada alumno, correspondiente a una serie de cuestiones abiertas. Esta evaluación, en la cual podría usarse el ordenador y se dio libertad para trabajar con cualquiera de los programas citados, correspondió al análisis de un nuevo fichero con 60 datos reales. El fichero recogía datos de una clase de educación física, con las variables sexo, práctica deporte habitualmente, pulsaciones por minuto en reposo, tiempo en recorrer 30 metros tras 30 flexiones en dos fechas distintas (septiembre y diciembre), peso, altura y la edad. A los alumnos se les dio una descripción del fichero, los datos impresos y una hoja de codificación del mismo.

Los autores analizan la forma en que los estudiantes usan el ordenador para resolver cuatro cuestiones de asociación. Como en cada ejercicio se puede solucionar por varios comandos distintos, nos muestran una tabla indicando las posibles soluciones correctas para cada uno de ellos.

Se toma el *intento* de solución como unidad de análisis, obteniendo 120 unidades de análisis para las cuatro cuestiones, es decir, 1,66 intentos de resolver por cada pregunta y alumno, por lo que hay alumnos que resuelven el problema por varios

caminos. Más adelante para cada unidad de análisis se estudian las siguientes variables: programa empleado en la resolución del problema y argumento en que el alumno basa su respuesta.

A continuación se analizan las estrategias correctas e incorrectas ofrecidas por los estudiantes. Las correctas encontradas se clasifican de la siguiente forma:

- *Frecuencias relativas*: Cuando los juicios de asociación se basan en proporciones o porcentajes en las distribuciones univariantes, condicionales o marginales de una tabla de contingencia.
- *Percentiles o rangos de percentiles*: Al basarse en frecuencias acumuladas ascendentes, descendentes, o en los estadísticos de orden (menos la mediana), en caso de datos ordinales o numéricos.
- *Coefficiente de correlación*: Valor absoluto y signo, en una tabla 2x2 u ordinal o bien coeficiente de contingencia en una tabla cualquiera.
- *Otros argumentos*: Gráficos que muestren la asociación, cálculos adicionales, etc.

En las dos primeras preguntas que aparece una variable dicotómica, los principales argumentos que se usan correctamente son los relacionados con frecuencias. Se deduce que la asociación de una variable con otra dicotómica se interpreta principalmente como diferencia en las distribuciones de frecuencias condicionales. Las otras preguntas se refieren a variables cuantitativas por lo que no las comentaremos.

Entre las argumentaciones incorrectas, la principal en la tabla de contingencia fue la interpretación incorrecta de las distribuciones de frecuencias condicionales. Siendo debido a un error de identificación de porcentajes y no darse cuenta de que la suma de porcentajes ha de ser igual a 100. Otro error es el uso de frecuencias absolutas: en las distribuciones univariantes, condicionales o marginales de una tabla de contingencia.

Además la existencia de otras variables parece que excluye la idea de dependencia. Hay quien piensa que la dependencia debe inducir un cambio de las frecuencias relativas de casos por encima y debajo de la media. Se ven casos que los alumnos encuentran respuestas distintas, sin darse cuenta que la respuesta es única. Finalmente se ve presencia de sistemas de creencias arraigadas y la no identificación del problema como de índole estadístico.

2.3.4.3. EVOLUCIÓN DE LAS CONCEPCIONES

Batanero, Estepa y Godino (1997) analizan la evolución de las concepciones de los alumnos tras la enseñanza. La mayor parte de los estudiantes de la muestra superaron la concepción determinista de la asociación, y la concepción local también se eliminó, al comprender la importancia de tener en cuenta los datos completos. La mayoría de los estudiantes usaron al finalizar la enseñanza las distribuciones condicionales, dejando los procedimientos aditivos.

Aparece alguna mejora en la concepción unidireccional, aunque todavía algunos estudiantes seguían considerando la independencia en la asociación inversa. En general no hubo mejora respecto a la concepción causal de la relación, ya que los estudiantes no se dieron cuenta que una asociación fuerte no es suficiente para llegar a una conclusión sobre la causa y el efecto. El estudio teórico de las relaciones entre la correlación y la causalidad durante la enseñanza fue insuficiente para cambiar esta concepción los estudiantes. Por consiguiente, creen los autores que hay una necesidad de encontrar las nuevas actividades prácticas que ayuden a los estudiantes a reflexionar sobre este tema.

2.3.4.4. ACTOS DE COMPRENSIÓN DE LA ASOCIACIÓN

Batanero, Estepa y Godino (1997) analizaron, mediante grabaciones, entrevistas y análisis de las tareas escritas y resueltas por ordenador, el proceso de enseñanza para dos estudiantes, con objeto de describir actos de comprensión de algunos elementos de significado de la asociación (Godino y Batanero, 1998). Interpretan los “actos de comprensión” según la teoría de Sierpiska (1994), descubriendo pruebas en los datos recogidos de que el conocimiento de los estudiantes de estos elementos de los significados parecen desarrollarse en unos momentos específicos durante el proceso. Estos elementos, algunos de los cuáles se refieren a la correlación numérica, fueron los siguientes:

1. *Para estudiar la asociación entre dos variables, la comparación de dos o más muestras tiene que ser hecha en términos de las frecuencias relativas* (Konold et al., 1997). En la primera sesión los estudiantes compararon las frecuencias absolutas de las mismas variables en dos muestras. Aunque se comentó sobre este error al final de esa sesión, el mismo procedimiento incorrecto aparecía otra vez en las sesiones 2, 3 y 5. Después, los estudiantes aparentemente superaron esta dificultad.
2. *La distribución completa en las diferentes muestras deben ser usadas para evaluar las diferencias en el estudio de la asociación.* Encontrar las diferencias locales no es

suficiente, puesto que la asociación debe ser deducida de los datos completos (Estepa, 1993). A pesar de esto, los estudiantes empezaron a solucionar los problemas comparando valores aislados en las dos muestras. Por ejemplo, solamente compararon los valores con las frecuencias máximas y mínimas para ambas muestras en la primera sesión. Aunque estas diferencias señalaban a una asociación posible, no eran suficientes para cuantificar su intensidad. Esta dificultad reapareció en las Sesiones 2 y 3 y desapareció definitivamente.

3. *De la misma frecuencia absoluta en una celda de la tabla de contingencia pueden ser calculadas dos frecuencias relativas condicionales diferentes, dependiendo de cuál es la variable condicionada.* El papel de la condición y lo que se condiciona en la frecuencia relativa condicional no es intercambiable. Falk (1986) y otros autores han señalado que los estudiantes tienen dificultades en la interpretación de las probabilidades condicionales, porque no discriminan entre la probabilidades $P(A/B)$ y $P(B/A)$. Muchos participantes en el estudio mostraron una confusión similar haciendo referencia a las respectivas frecuencias condicionales en la prueba preliminar y durante todas las sesiones experimentales. Esta confusión fue notada en los estudiantes observados durante sesión 5, aunque lo solucionaron con ayuda. Ellos no mostraron esta confusión durante el resto de las sesiones.
4. *Dos variables son independientes si la distribución de una de estas variables no cambia cuando la condicionamos por valores de la otra variable* (Nortes, 1993). Hasta sesión 5, los estudiantes no descubrieron que una condición para la independencia era invariable respecto de la distribución de frecuencia condicional cuando cambia el valor de la variable condicional.
5. *La decisión sobre qué tamaño de las diferencias deber ser considerado para admitir la existencia de la asociación es, hasta cierto punto, subjetiva.* Es difícil obtener la relación perfecta o la independencia. El problema de la relación debe ser planteado en términos de intensidad en vez de en términos de existencia (Estepa, 1993). Aunque los estudiantes no habían estudiado constantes de hipótesis, en la sesión 5 descubrieron que determinar un juicio de asociación implica tomar una decisión sobre si atribuir las diferencias pequeñas a las fluctuaciones de muestra o a la relación verdadera entre las variables. También se dieron cuenta de que hay diferentes grados de asociación, desde independencia perfecta a relación funcional.
6. *Cuando se estudia la asociación ambas variables juegan un papel simétrico.* Sin embargo, cuando investigamos la regresión el papel jugado por las variables no es

simétrico. El hecho es que esa correlación ignora la distinción entre explicativas y variables de respuesta, mientras en la regresión esta diferencia es esencial (Moore, 1997), causa mucha confusión para los estudiantes. Cuando tuvieron que seleccionar la variable explicativa para encontrar sentido a la línea de regresión, en las sesiones 5, 6 y 7, no supieron qué variable escoger. Por ejemplo, cuando calculaban la línea de regresión entre la altura y el peso, los estudiantes fueron confundidos por el hecho de que había una dependencia mutua de las dos variables. Una gran cantidad de discusión siguió, ya que los estudiantes no eran capaces de solucionar esta confusión. El profesor no notaba el problema y definitivamente, los estudiantes calculaban las líneas de regresión escogiendo la variable explicativa al azar. Al final del período de enseñanza estos estudiantes no habían descubierto que dos líneas de regresión diferentes pueden ser calculadas.

7. *Una correlación positiva señala a una asociación directa entre las variables.* Aunque, en la sesión 6, estudiantes podían interpretar el tamaño del coeficiente de correlación, no hablaron del tipo de la asociación (directa o inversa). Al final de la sesión, notaban que cuando el coeficiente de correlación es positivo, y hay una relación lineal, las variables son asociadas positivamente y por encima del promedio. Sin embargo, no usaron el termino "Asociación directa" explícitamente.
8. *Una correlación negativa señala a una asociación inversa entre las variables.* Cuando, en la sesión 6, los estudiantes encontraron un coeficiente de correlación negativo por primera vez, se extrañaron tanto que preguntaron a su profesor si esto era posible. También tenían problema cuando comparaban dos coeficientes de correlación negativos. Aunque, con la ayuda del profesor, observaron que un coeficiente de correlación negativo correspondía a una pendiente negativa de la línea de regresión y que esto quiso decir que el valor de y disminuyó cuando el valor de x aumentó, no usaron el termino "asociación inversa" explícitamente, ni diferenciaron entre los dos tipos de la relación al final de su aprendizaje.
9. *El valor total del coeficiente de correlación indica la intensidad de la asociación.* Aunque los estudiantes relacionaron el valor total del coeficiente de correlación con la intensidad de la relación, no relacionaron esta idea con la dispersión de los datos.

Las conclusiones de la investigación citada indican que un proceso de enseñanza de nociones estadísticas basado en la resolución de problemas con datos reales de

análisis de datos, disponiendo de soporte informático, es insuficiente para que los estudiantes adquieran todo el sentido y verdadero alcance de la asociación e independencia. El diseño de situaciones didácticas debe tener en consideración las distintas variables didácticas, y adjudicar un papel importante al profesor gestionando las situaciones. En todo caso, los estudiantes eran futuros profesores de educación primaria, con poca base estadística previa. En nuestra experiencia de enseñanza (Capítulo 4) nos centramos en estudiantes de psicología que ya han tenido un curso previo de estadística, incluida la inferencia y analizamos exclusivamente (también con más detalle) la comprensión de la asociación en tablas de contingencia.

2.4. CÁLCULO DE PROBABILIDADES A PARTIR DE TABLAS DE CONTINGENCIA

Un segundo tipo de investigaciones es el relacionado con cálculo de probabilidades a partir de datos en una tabla de contingencia. Los objetos matemáticos presentes en las tablas de contingencia pueden ser confundidos por los alumnos al resolver estos problemas. Además de las frecuencias dobles, condicionales y marginales que aparecen en la tabla, obtenemos diversas probabilidades que se deducen de cada una. Estas probabilidades pueden ser probabilidad conjunta ($P(A \cap B) = \frac{a}{a+b+c+d}$),

probabilidad condicional respecto a su fila ($P(A|B) = \frac{a}{a+b}$), probabilidad condicional

respecto a su columna ($P(B|A) = \frac{a}{a+c}$), y probabilidad simple de filas

($P(B) = \frac{a+b}{a+b+c+d}$), o de columnas ($P(A) = \frac{a+c}{a+b+c+d}$).

El cálculo de estas probabilidades trae consigo muchos problemas, como la *falacia de la condicional transpuesta*, que consiste en no distinguir entre las dos direcciones de la probabilidad condicionada (Falk, 1986), interpretar incorrectamente el conjunto “y”, confundiendo la probabilidad conjunta y la probabilidad condicional (Einhorn y Hogarth, 1986), o interpretar la intersección como condicionamiento (Ojeda, 1995).

En lo que sigue resumimos los trabajos relacionados con el tema, que han sido llevado a cabo con diferentes tipos de estudiantes: (a) estudiantes de psicología, (b) futuros profesores; y (c) estudiantes de secundaria.

2.4.1. INVESTIGACIÓN CON ESTUDIANTES DE PSICOLOGÍA

El trabajo con este tipo de estudiantes es iniciado por Díaz y de la Fuente (2005a), quienes trabajan con una muestra de 154 estudiantes de psicología, usando tablas de contingencia 2x2, encontrando muchos posibles errores en el cálculo de probabilidades a partir de estas tablas. Las autoras sugieren que estos serán posiblemente más frecuentes todavía al interpretar tablas más complejas. En lugar de pedir a los alumnos que realicen un estudio de la asociación, las autoras evalúan si los estudiantes identifican y calculan la probabilidad simple, condicional y conjunta a partir de los datos de la tabla.

En la investigación de Díaz y de la Fuente (2005a), analizan las respuestas a cuatro apartados de una pregunta de un cuestionario más amplio que se pasó después de impartir el tema de probabilidad y motivando a los alumnos puntuándoles el número de respuestas correctas para la nota final de una asignatura.

En esta investigación continúa categorizando las respuestas, y mostrando los porcentajes de respuestas correctas en cada uno de los cuatro apartados del ejercicio. Se aprecia que el 90% de los participantes responden correctamente al cálculo de probabilidades simples, pero más del 40% fallan al tratarse de calcular probabilidades compuestas y condicionales. Posteriormente clasifican las distintas respuestas obtenidas en una tabla, y comenta los siguientes errores:

- *Confundir un suceso con su complementario.* En lugar de calcular la probabilidad pedida se calcula la del complementario, tanto en probabilidades simples como en probabilidades condicionadas, o sea de $P(A/\bar{B})$, en lugar de $P(A/B)$.
- *Confundir probabilidades con casos posibles,* dando el valor de la celda correspondiente a frecuencias absolutas, como si fueran probabilidades de intersección.
- *Obtener probabilidades mayores que la unidad,* calculando las probabilidades dando la vuelta o invirtiendo la fórmula de Laplace.
- *Confundir la unión con la intersección de sucesos.* Aplicando a la palabra “y” del enunciado un significado de unión, frecuente en el lenguaje cotidiano, pero no igual en este caso con el significado matemático.
- *Dar por hecho la independencia de los datos.* Estos alumnos calculan la probabilidad de la intersección multiplicando las dos probabilidades de los dos sucesos.

- Al pedir la probabilidad condicionada dar como respuesta $\frac{P(A)}{P(B)}$. Estando presente un conflicto en confundir el numerador del cociente de la probabilidad condicionada.

Se hace una nueva clasificación de los errores atendiendo los diferentes conflictos semióticos entre objetos matemáticos presentes en la tarea, encontrándose 262 conflictos en las cuatro tareas y total de muestra, esto es una media de 1,7 conflictos por alumno encontrados, ya que una respuesta puede traer más de un conflicto semiótico. Se aprecia como errores principales confundir una probabilidad condicional con una conjunta o confundir la probabilidad condicional con su inversa. Muchos estudiantes no distinguen un suceso con su complementario, tienen errores de cálculo o no encuentran los datos en el enunciado. Algunos errores menos comunes son confundir frecuencia con probabilidad o poner al contrario la fórmula dando probabilidades mayores de 1.

2.4.2. TRABAJOS CON FUTUROS PROFESORES

Recientemente ha crecido el contenido de probabilidad en la Enseñanza Primaria, (MEC, 2006), dentro del bloque 4, Tratamiento de la información, azar y probabilidad. Ello implica que los futuros profesores deben adquirir competencias para afrontar análisis de experiencias aleatorias o utilizar expresiones relacionadas con la probabilidad, estimando o calculando probabilidades sencillas. Actualmente no se ve el acercamiento a dichas competencias en la formación de profesores a nivel internacional (Franklin y Mewborn, 2006). El *Joint ICMI/IASE Study Teaching Statistics in School Mathematics* (Batanero, Burrill, Reading y Rossman, 2008) trató de contribuir en este aspecto.

La comprensión de las tablas de contingencia por parte de futuros profesores apenas ha sido estudiada, a pesar de su importancia y complejidad. Una excepción fue el experimento de enseñanza realizado por Estepa (1993) con una muestra de futuros profesores y que hemos descrito en las secciones anteriores. Otras autoras que tratan el tema son Estrada y Díaz (2006), quienes evalúan las dificultades de futuros profesores de educación primaria en problemas sencillos de la vida cotidiana. Las autoras analizan las respuestas de una muestra de 65 futuros profesores de la Universidad de Lleida cuyo

plan de estudios contempla una asignatura troncal de Matemáticas y su didáctica, con una parte sobre Estadística y Probabilidad, y una optativa de Estadística.

Les plantean el mismo problema usado por Díaz y de la Fuente (2005a), que incluye el cálculo de una probabilidad simple, otra compuesta y otra condicional a partir de datos dados en una tabla de contingencia. El porcentaje de respuestas correctas en todo lo que se preguntó superó el 50%, y superior al 75% en el caso de la probabilidad simple, pero de todos modos es pequeño, al tratarse de preguntas muy simples que el futuro profesor puede encontrar en la vida diaria o profesional. Las autoras encuentran en sus estudiantes conflictos como confusión entre probabilidades simples compuestas y condicionales, así como confusión entre un suceso y su complementario, probabilidad con casos favorables o suponer independencia en los datos.

Otro trabajo en el que se trata este tema es en Contreras, Estrada, Díaz y Batanero (2010), donde tratan de completar el trabajo de Estrada y Díaz (2006), centrándose en 69 futuros profesores de Educación Primaria de la Universidad de Granada sin una formación específica en estadística. Además simplifican el enunciado del problema planteado, evitando el uso de negaciones y desigualdades, para hacerlo más asequible a los participantes. Los alumnos participantes en el estudio habían realizado dos prácticas dedicadas a los temas de estadística y probabilidad, y además en el año anterior al que se tomaron los datos, los alumnos habían cursado la asignatura Matemáticas y su Didáctica, con parte de estadística y azar.

Recogidos los datos, en esta investigación se realizó un análisis semiótico de las respuestas abiertas siguiendo el método propuesto por Font, Godino y D'Amore (2007), tratando de identificar los conflictos semióticos de los estudiantes. En el problema que se plantea en esta investigación, se clasifican 700 niños imaginarios en función de dos variables dicotómicas (le gusta o no el tenis; es chico o chica), representándolo en el enunciado en una tabla de contingencia. Mostrándose las cuatro celdas centrales (frecuencias absolutas), el total de sujetos y las frecuencias absolutas marginales de cada valor de cada variable. Las probabilidades pedidas en los tres apartados del problema son la probabilidad simple $P(A)$, la conjunta $P(A \cap B)$ y la probabilidad condicional $P(A/B)$, en el caso de equiprobabilidad. El objetivo es identificar conflictos semióticos y compararlos con los descritos por Estrada y Díaz (2006).

A pesar de la mayor facilidad del ítem, en esta investigación los resultados son peores que en la anterior, lo que explicamos porque los estudiantes de Estrada y Díaz seguían un curso optativo de Didáctica de la Estadística. Muchos estudiantes no

aportaron solución, lo que indica falta de competencia en la lectura de la tabla doble. Algunos conflictos semióticos encontrados fueron:

- Confunden la probabilidad condicional $P(A/B)$ con $P(B/A)$.
- Confusiones entre la probabilidad condicional y conjunta e incluso unos pocos casos que confunden la simple y la condicional o la conjunta.
- Confunden probabilidad con casos posibles o favorables.
- Se confunde un suceso con otro.
- Se calcula la probabilidad conjunta como producto de las probabilidades de los sucesos, suponiendo independencia en las variables, lo que implica una falta de percepción de la asociación entre ellas.
- Otros identifican correctamente las probabilidades pero confunden las fórmulas.
- Confunden casos favorables con posibles.
- Dan probabilidades mayores que 1.

Contreras, Batanero, Díaz y Fernandes (2011) completan el anterior estudio sobre dificultades en el cálculo de probabilidades (simple, conjunta y condicional) de una tabla de contingencia en una muestra de 183 futuros profesores de educación primaria de la Universidad de Granada, con el mismo problema usado en la investigación anterior, aunque completando el análisis.

Los resultados del estudio se presentan con una clasificación basada en la exactitud de la respuesta y el tipo de errores (en caso de error), siendo las categorías: respuestas básicamente correctas, confundir probabilidades, confundir sucesos, confundir fórmulas, otros errores o no finaliza el problema. La investigación se completa con dos preguntas adicionales. En una de ellas, se pidió a los participantes que identificasen el contenido de probabilidad necesario para resolver la tarea. Se incluyeron las siguientes categorías de objetos:

- *Problemas*: Se esperaba que los estudiantes identificasen los tres problemas diferentes específicos en la tarea: Un problema de probabilidad simple, un problema de probabilidad compuesta, y un problema de probabilidad condicional.
- *Lenguaje*: En el enunciado de la tarea aparece lenguaje matemático verbal, numérico y tabular; dependiendo de la solución, algunos estudiantes también hacen uso del lenguaje simbólico y gráfico.
- *Conceptos*: podemos identificar implícito en la tarea, los conceptos de experimento

aleatorios (selección de un estudiante de la escuela al azar); suceso simple y compuesto; muestra, casos favorables y desfavorables; probabilidad simple, probabilidad compuesta y probabilidad condicional, fracción, proporción y porcentaje, frecuencia y porcentaje, número entero, operaciones con números enteros (división).

- *Propiedades* (o relaciones entre conceptos). Algunas propiedades implícitas en esta tarea son: Los axiomas de probabilidad, la relación entre la probabilidad de un suceso y la de su contrario, el hecho de que el espacio muestral se reduce al calcular la probabilidad condicional, la equivalencia de dos fracciones al dividir los dos términos de la fracción por el mismo número, la regla de Laplace, la relación entre el tamaño de la muestra total y los totales en filas o en columnas, la relación entre las frecuencias dobles, marginales y condicionales.
- *Procedimientos* (o algoritmos). Procedimientos que se pueden utilizar en la solución de estas tareas incluyen hacer operaciones numéricas, tales como división o adición, simplificación de fracciones, lectura de una tabla, transformación de una probabilidad en porcentaje, aplicación de las fórmulas para el cálculo de probabilidades simples, compuestas y condicionales, y cálculo de porcentajes o proporciones.
- *Argumentos*. El principal tipo correcto de argumento utilizado para resolver la tarea es el argumento deductivo, que fue identificado por muchos estudiantes.

Muchos estudiantes fueron capaces de identificar y clasificar correctamente algunos de los objetos matemáticos en el problema, aunque, en general, el número de objetos identificados fue muy pequeño, y una proporción importante de estudiantes no fueron capaz de dar ejemplos en algunas categorías. Otros ejemplos proporcionados por los estudiantes fueron considerados incorrectos, debido a algunas de las siguientes razones:

- Algunas de las respuestas eran demasiado imprecisas, por ejemplo, se responde que un problema matemático es "responder las preguntas que aparecen después de la tabla de datos".
- Algunos estudiantes confunden los diferentes tipos de objetos matemáticos, por ejemplo, consideran los procedimientos "la interpretación de la tabla" o "realizar una división" como conceptos.

- Otros estudiantes incluyeron en sus respuestas algunos de los objetos matemáticos que no eran necesarios para resolver la tarea, como, por ejemplo, "mediana, moda, desviación estándar".

Los resultados sugieren que la identificación de los objetos matemáticos implícitos en la tarea no fue fácil para los participantes en la muestra. Los elementos más identificados por los futuros profesores fueron los conceptos (2-3 conceptos correctamente identificados por participante), procedimientos y lenguaje (1-2 correctamente identificados).

Los resultados de estos trabajos son motivo de preocupación, pues las tablas de doble entrada y su interpretación aparecen como contenido a enseñar en los decretos de Educación Primaria. Para poder seleccionar adecuadamente tareas para la enseñanza, y poder resolver las dificultades de los niños, los profesores, en primer lugar requieren un conocimiento matemático sólido por parte del profesor (Ball, Lubienski y Mewborn, 2001). El futuro educador debe dominar el tema para transmitirlo en la enseñanza, por tanto en estas investigaciones se ve la necesidad de mejorar la educación estadística que los futuros profesores reciben durante su formación en las facultades de educación.

2.4.3. OTROS TRABAJOS

Siguiendo esta línea de investigación, Lonjedo y Huerta (Huerta y Lonjedo 2003; Lonjedo, 2003; Lonjedo y Huerta, 2005) describen algunas variables que influyen sobre la resolución de problemas de probabilidad condicional dados en una tabla de contingencia por estudiantes con y sin nociones previas de los contenidos. Los autores analizan si la resolución de los problemas puede hacerse utilizando el razonamiento numérico y por tanto los estudiantes no necesitan utilizar las relaciones entre probabilidades para resolver el problema. Esto ocurre sobre todo cuando los problemas no son interpretados como probabilidades y a causa de ello no se usan las propiedades de la probabilidad para obtener la solución del problema. Es sólo al final del proceso de resolución del problema cuando los estudiantes responden a la pregunta del problema en términos de probabilidad.

Lonjedo y Huerta concluyen que hay factores que afectan al éxito en la resolución del problema, que no necesariamente son el conocimiento de las relaciones entre probabilidades. En su estudio clasifican los problemas atendiendo a la naturaleza de los datos en el texto del problema, distinguiendo entre datos presentados en términos

de razón o datos expresados en una combinación de estos formatos. Los autores presentaron a los alumnos seis problemas de probabilidad condicional en el que la estructura de datos no varía, pero sí su naturaleza y el contexto. Los resultados mostraban que el éxito en la resolución de los problemas no dependía del uso correcto de unas determinadas fórmulas. Estudiantes capacitados para ello, no siempre usaban los datos interpretados como probabilidades. Por otra parte, estudiantes que no conocen dichas fórmulas resuelven los problemas, por lo que el éxito no depende de ellas sino del uso correcto del razonamiento aritmético aplicado a unos datos no interpretados como probabilidades, sino como razones o proporciones.

Uso de frecuencias naturales

Gigerenzer (1993) sugiere que la dificultad en la resolución de problemas de probabilidad condicional, en particular, los referidos al teorema de Bayes desaparece cuando las preguntas se plantean en términos de frecuencias. Llama a este formato *frecuencias naturales* porque se asemeja más a la forma en que recogemos información de las frecuencias de sucesos aleatorios en una situación de muestreo natural a lo largo de nuestra experiencia. Cuando la información se ofrece en términos de frecuencia, el cálculo de la probabilidad condicional es más natural; por ejemplo, en el cálculo de probabilidad a posteriori el sujeto no tiene que aplicar toda la complejidad del teorema de Bayes, sino sólo tener en cuenta los casos favorables y posibles, de modo que el problema se transforma en un problema simple de probabilidad.

Resultados similares se han obtenido al plantear otros problemas probabilísticos en términos frecuenciales, por lo que Gigerenzer recomienda, cuando sea posible, cambiar a frecuencias el formato de las preguntas. Esta idea es retomada por Martignon y Wassner (2006), junto con un diagrama que refuerza la comprensión de los pasos de resolución para proponer una enseñanza más simple de este teorema.

2.5. COMPRENSIÓN DEL CONTRASTE DE HIPÓTESIS

Es frecuente al analizar las tablas de contingencia para el análisis de asociación entre variables o de homogeneidad de muestras utilizar el contraste Chi-cuadrado. (Lipschutz y Schiller, 1999; Ruiz-Maya y Martín, 2002). En el Capítulo 1 observamos como aparecían estos contrastes entre los procedimientos del manejo de tablas de contingencia y en la enseñanza que se realizará se incluye este contraste en el Tema 3.

Será entonces importante tener en cuenta las posibles dificultades de los estudiantes para comprender los pasos necesarios en un contraste de hipótesis.

Los estudiantes conocen ya este tema, pues en la asignatura “Descripción y exploración de datos en Psicología”, del grado de Psicología en la Universidad de Granada, entre sus contenidos aparece “*generalización e inferencia y formulación y contraste de hipótesis: significación estadística frente a relevancia práctica*”, que se incluye debido a la importancia que tienen los contrastes de hipótesis para el ámbito psicológico. Otro motivo para estudiar este contenido son los resultados de la investigación previa, que nos indican la gran dificultad de comprensión y aplicación del contraste de hipótesis, no sólo entre estudiantes, sino incluso en los investigadores en diversas ciencias, que lo usan en sus investigaciones, aunque no siempre en forma adecuada (Morrison y Henkel, 1970; Harlow, Mulaik y Steiger, 1997; Borges, San Luis, Sánchez y Cañadas, 2001).

Estas malas prácticas han sido atribuidas a la dificultad filosófica y psicológica del tema, que se añade a la complejidad matemática del contraste de hipótesis, en que el estudiante ha de comprender y relacionar conceptos como parámetro y estadístico, población y muestra, distribución muestral, región crítica y de aceptación, error tipo I y II y sus posibilidades y p -valor así como la lógica subyacente (Harradine, Batanero y Rossman, 2011).

Las dificultades filosóficas son debidas a la aparente similitud del contraste de hipótesis con la demostración por reducción al absurdo y al deseo de los investigadores de justificar, mediante la probabilidad, la validez del razonamiento inductivo (Rivadulla, 1991; Batanero, 2000). Popper (1967) sugirió poner a prueba las hipótesis científicas, mediante experimentos y comparar los patrones deducidos de la teoría con los datos obtenidos. Pero, mientras que en una demostración o “prueba” matemática el razonamiento es deductivo y tiene, por tanto, una validez lógica, esto no ocurre en el contraste de hipótesis.

Las dificultades psicológicas son debidas al deseo de control de la incertidumbre asociada a la toma de decisión en un contraste (Gigerenzer, 1993). Pero esta incertidumbre, que se debe al proceso de muestreo, nunca puede obviarse. Precisamente los conceptos de error tipo I y II y potencia, tratan de cuantificar en forma de probabilidad los riesgos inherentes en el contraste de hipótesis.

Como didactas de la matemática, nosotros tratamos de explicar una parte de estos errores por deficiencias del sistema de enseñanza. Por ejemplo, Brewer (1986)

informa de una investigación donde analiza la forma en que los libros de texto presentan el contraste de hipótesis e indica que los libros de texto y la literatura estadística utilizada en las ciencias del comportamiento contiene con frecuencia errores, que podrían ser los principales causantes de los errores de aprendizaje y de las aplicaciones incorrectas de la estadística inferencial. Posteriormente Brewer (1989) propone el uso de analogías para la enseñanza de la lógica del proceso de contraste de hipótesis. También Vallecillos (1998) hace algunas propuestas generales para mejorar la enseñanza de la estadística inferencial en los distintos niveles de enseñanza y Harradine, Batanero y Rossman (2011) sugieren una serie de pasos para desarrollar el razonamiento inferencial, primero de manera informal, y luego formal en los estudiantes.

Sin intentar detallar esta problemática, a continuación nos centramos en las investigaciones que describen las dificultades asociadas a la interpretación de los conceptos matemáticos que intervienen en un contraste de hipótesis.

Investigaciones sobre el planteamiento de las hipótesis

Son muchas las investigaciones sobre contrastes de hipótesis que muestran que los alumnos confunden conceptos necesarios en los contrastes de hipótesis (Falk, 1986; Vallecillos y Batanero, 1997).

El primer paso en un contraste de hipótesis es plantear las hipótesis, lo cuál no es siempre sencillo para los estudiantes. En primer lugar, Chow (1996) analiza la diferencia entre diversos tipos de hipótesis: sustantiva o teórica (que se refiere al marco teórico del estudio y se suele aceptar sin someterla a prueba); de investigación (la que interesa al investigador, que suele ser muy amplia); la experimental (una restricción de la hipótesis de investigación, generalmente se refiere a una variable que el investigador puede analizar); finalmente llegamos a las hipótesis estadística.

Dentro de ellas hemos de diferenciar la hipótesis alternativa y la hipótesis nula, que se refieren a un parámetro u otra característica de la distribución de una variable en una población de sujetos. La hipótesis alternativa es la que el investigador desea contrastar; pero en estadística se trabaja con la hipótesis nula o de no efecto. Esta hipótesis es la que nos sirve para fijar la distribución muestral del estadístico de contraste que se utiliza para contrastarla (Batanero, 2000).

Es importante diferenciar estos niveles de hipótesis, pues algunos investigadores cuando se encuentra un resultado significativo piensan que el resultado se puede

extrapolar a la hipótesis experimental o incluso a la de investigación (Chow, 1996). Vallecillos (1994) entre sus resultados con una muestra de 436 estudiantes de diferentes especialidades (empresariales, estadística, ingeniería, medicina y psicología) encuentra que el 13% de alumnos que confundieron la hipótesis nula con la alternativa.

Respecto a la forma de plantear las hipótesis. Vallecillos (1994) encuentra alrededor de un 20% de sujetos de su muestra, que no sabían si la hipótesis del contraste se refiere al parámetro de la población o al estadístico muestral (utilizando el estadístico muestral para plantear sus hipótesis). Ello es debido a que los estudiantes confunden parámetro y estadístico, por ejemplo, la media de la población (que es constante pero desconocida) con la media de la muestra (que se conoce para la muestra particular, pero es una variable aleatoria en el conjunto de todas las muestra del mismo tamaño tomadas de la población) (Schuyten, 1991; Vallecillos y Batanero, 1997; Castro Sotos et al., 2007).

En la investigación de Olivo (2008) en uno de sus ítems sobre intervalos de confianza el 29% de los estudiantes confunden la media poblacional con la media muestral y que el 25,8% de estudiantes confunden estadístico y parámetro. Comprender el estadístico como variable aleatoria supone comprender la distribución muestral, pero Harradine, Batanero y Rossman (2011), indican que el concepto de distribución muestral es mucho más abstracto que los de distribución de una población o de una muestra, lo que puede causar dificultad al expresar las hipótesis. Schuyten (1991) indica que los estudiantes confunden los diversos planos en que se aplica un mismo concepto en inferencia, como sería en el caso del concepto de media muestral y de media poblacional.

En la investigación de Vallecillos, se determina tres concepciones principales sobre el papel que desempeña la hipótesis nula en un contraste:

- *Hipótesis nula como hipótesis a demostrar.* Los alumnos manifiestan una concepción de la hipótesis nula que se asemeja más a un enunciado a demostrar que a una opción a elegir entre las dos posibles en un problema de decisión.
- *Hipótesis nula referida indistintamente a la población o a la muestra.* Las hipótesis que se plantean en un contraste se refieren siempre a una población determinada, los datos con los que se trabaja pertenecen a una muestra de la misma. Los estudiantes plantean hipótesis que se refieren a la muestra en lugar de la población.

- *Hipótesis nula referida a una única población o a un único parámetro.* Los alumnos manifiestan una concepción muy restrictiva de los tipos de hipótesis admisibles.

Vera, Díaz y Batanero (2011) con una muestra de 224 estudiantes de Psicología después de realizar un curso de inferencia, clasifican los errores del planteamiento de hipótesis en dos categorías, según utilicen el parámetro de la población o el estadístico muestral. Entre los que plantean las hipótesis usando el parámetro (media de la población), encontraron errores, como intercambiar el contraste unilateral y bilateral, plantear una hipótesis alternativa puntual, y por tanto no se cubre el espacio paramétrico o plantear la hipótesis nula y alternativa mediante intervalos no disjuntos, que, además no cubren el espacio paramétrico. Por otro lado, entre las respuestas incorrectas que usan explícitamente la media muestral encuentran los siguientes errores: Plantear como hipótesis la igualdad entre media muestral y poblacional, confundir contraste unilateral y bilateral. Explican estos errores por medio de los siguientes conflictos semióticos:

- Conflictos de reconocimiento del campo de problemas: confusión entre contraste unilateral y bilateral; y reconocer incorrectamente un problema de análisis de varianza como contraste de comparación de dos medias.
- Conflictos de reconocimiento de propiedades de las hipótesis: conflicto al plantear una hipótesis alternativa puntual; conflicto al definir hipótesis no complementarias; plantear hipótesis no disjuntas; y plantear las hipótesis en función del estadístico muestral.
- Conflictos de confusión de otros objetos que intervienen en el contraste: confusión entre el estadístico muestral y el parámetro correspondiente a la media poblacional; y confundir la muestra con la población.
- Otros: Confundir la variable que está siendo contrastada.

Comprensión del nivel de significación y el valor p

La mayoría de las investigaciones sobre el contraste de hipótesis se centran en la comprensión de estos conceptos e indican su dificultad. El nivel de significación α , se define como la probabilidad de rechazar la hipótesis nula en caso de que sea cierta, pero los estudiantes e investigadores intercambian los dos términos de la probabilidad condicional, es decir, interpretan el nivel de significación como la probabilidad de que la hipótesis nula sea cierta si hemos tomado la decisión de rechazarla.

Birnbaum (1982), por ejemplo, informó que sus estudiantes encontraban razonable la siguiente definición: "Un nivel de significación del 5% indica que, en promedio, 5 de cada 100 veces que rechazamos la hipótesis nula estaremos equivocados". Resultados semejantes se describen en otros estudios (Bakan, 1966; Birnbaum, 1982; Vallecillos, 1994; Krauss y Wassner, 2002). El mismo error se produce cuando se interpreta el valor p (probabilidad de obtener el valor dado del estadístico u otro más extremo cuando la hipótesis nula es cierta). También se han descrito los mismos errores en profesores de metodología e investigadores en Psicología (Lecoutre, 1999, 2006; Lecoutre, Lecoutre y Poitevineau, 2001; Haller y Graus, 2002)..

Para explicar estos errores, Falk (1986) recuerda que el nivel de significación viene dado por una probabilidad condicional, pero la ambigüedad verbal en la expresión "error tipo I", que es una expresión no condicionada, puede provocar confusión, ya que se puede creer que se está trabajando con la probabilidad de un suceso simple. Menon (1993), en la misma línea, afirma que la expresión *error tipo I* lleva a la idea de un único suceso, tendiendo la gente a olvidar que trabajan con probabilidades condicionales. Otra explicación es la falta de diferenciación en ciertos sujetos entre la probabilidad condicional $P(A/B)$ y $P(B/A)$, conocida como *falacia de condicional traspuesta* o *confusión de la inversa* por Falk (1986).

En el trabajo de Vallecillos (1994) se encuentran una gran variedad de interpretaciones del nivel de significación, la mayoría incorrectas. Un error prevalente en todos los grupos de estudiantes fue el intercambio de los términos de la probabilidad condicional, incluso entre estudiantes que eran capaces de discriminar entre una probabilidad condicional y su inversa. La autora describe tres tipos principales concepciones incorrectas sobre el nivel de significación:

- *Nivel de significación como probabilidad condicional referida a alguna de las hipótesis.* Los alumnos conservan la idea de probabilidad condicional en la definición del nivel de significación, pero cambiando los sucesos condición y condicionado o ambos. También Díaz, Batanero y Wilhelmi (2008), indican que la interpretación incorrecta más frecuente sobre los conceptos del contraste de hipótesis es cambiar los términos de la probabilidad condicional en su definición, considerando el nivel de significación como la probabilidad de que la hipótesis nula sea cierta.
- *Nivel de significación como probabilidad simple de la hipótesis nula.* Los alumnos

interpretan el nivel de significación como probabilidad no condicional de la hipótesis nula, tanto en el caso de aceptación como de rechazo. La autora indica que no tiene sentido calcular la probabilidad de una hipótesis en el paradigma de la estadística frecuencial, pero sí en la estadística bayesiana.

- *Nivel de significación como probabilidad de error.* Las respuestas que se refieren a como la probabilidad de error, tanto si se asocia a la hipótesis nula como a la alternativa e indistintamente si se acepta o se rechaza cualquiera de ellas.

Williams (1997, 1998) analiza la comprensión del concepto de nivel de significación en su investigación con 18 estudiantes universitarios de un curso de estadística a los que propuso tareas de tipo conceptual y la realización de dos problemas para analizar el conocimiento de tipo procedimental. Entre sus conclusiones indica:

“el conocimiento de la mayoría de los estudiantes sobre el nivel de significación era limitado. Sin embargo, la definición correcta y la descripción precisa de su papel en el proceso de toma de decisión parecía estar asociado con una mejor aplicación del procedimiento” y “el estudio pone de manifiesto muchos problemas de los estudiantes asociados con el conocimiento del concepto” (Williams, 1998, p. 748).

Otros errores señalados por Batanero (2000) sobre el nivel de significación y el valor p son los siguientes:

- Pensar que el valor p es la probabilidad de que el resultado se deba al azar. Podemos ver que esta concepción es errónea ya que un resultado significativo puede ser debido a otros factores.
- Suponer que se conserva el nivel de significación cuando se realizan contrastes consecutivos en el mismo conjunto de datos, lo que produce el problema de las comparaciones múltiples. En ocasiones aplicamos un alto número de pruebas de significación a un mismo conjunto de datos, con lo que el nivel de significación global crece, lo que no es apreciado por los investigadores.
- El uso frecuente de niveles de significación .05 y .01 es cuestión de convenio y no se justifica por la teoría matemática. Sin embargo suele pensar que existe una justificación matemática para considerar un nivel de significación de 0,05 o de 0,01. Esto hace que investigaciones que son significativas al 0,05 y que no lo son al 0,06 no se publiquen (Skipper, Guenter y Nass, 1970).
- La interpretación incorrecta del nivel de significación aparece junto a una

interpretación incorrecta de los resultados significativos. Según Frías, Pascual y García (2000), la significación estadística por si sola no implica una significación práctica, ya que se encuentran datos significativos con un pequeño efecto experimental al tomar una muestra grande. Esto no suele ser comprendido (Lecoutre, 1999). Un resultado significativo implica para Fisher que los datos proporcionan evidencia en contra de la hipótesis nula, mientras que para Neyman y Pearson solo establece la frecuencia relativa de veces que rechazaríamos la hipótesis nula cierta a la larga (error tipo I). Pero Lecoutre (1999), en su investigación con 20 investigadores de Psicología y 25 profesionales estadísticos, informa de la confusión entre significación estadística y significación práctica (significación estadística y un efecto experimental).

Otras dificultades

En otras investigaciones se menciona que los alumnos no llegan a tomar la decisión adecuada, por fallo en la comprensión de la lógica del contraste (Birnbaum, 1982; Vallecillos, 1994; Lecoutre, Lecoutre y Poitevineau, 2001; Haller y Kraus, 2002; Lecoutre, 2006; Harradine, Batanero y Rossman, 2011).

En algunos casos, los sujetos confunden los valores que puede tomar una variable, con las distintas variables, al indicar que cada valor de una variable, es una variable diferente. White (1980) describe este error y supone que se debe a la dificultad que tienen algunos investigadores para identificar la población bajo estudio al aplicar la inferencia, coincidiendo también con Hagod (1970)

Schuyten (1991) indica que los estudiantes confunden los diversos planos en que se aplica un mismo concepto en inferencia, como sería en el caso del concepto de media muestral y de media poblacional.

Respecto a las condiciones de aplicabilidad de los contrastes de hipótesis, al aplicar un test Chi-cuadrado para estudiar la posible asociación entre dos variables en una tabla de contingencia, los supuestos podrían ser violados, bien por tener una frecuencia pequeña en cada casilla, bien por haber reagrupado “convenientemente” filas o columnas en la tabla para que aparezca una cierta asociación, o bien por realizar la prueba en tablas binarias cuando se tiene un conjunto de datos multivariante, en que una relación podría implicar más de una de dichas variables (Lipset, Trow y Coleman, 1970). Este punto podría no ser comprendido por los estudiantes.

En el contraste de homogeneidad o independencia se requiere la obtención de los

grados de libertad, Walker (1940) reflexionó sobre la dificultad del concepto de grados de libertad, que también se observó 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 . Todos estos puntos pudieran influir en la realización del contraste Chi-cuadrado y por tanto hemos de tenerlos en cuenta en nuestro trabajo.

2.6. CONCLUSIONES DEL ESTADO DE LA CUESTIÓN

En este capítulo hemos mostrado los antecedentes de nuestro trabajo, que se pueden dividir en tres tipos: (a) estudios relacionados con la estimación y juicios de asociación en tablas de contingencia; (b) investigaciones sobre cálculo de probabilidades en estas tablas, mostrando la gran relación que hay presente entre los conceptos de probabilidad y asociación y (c) investigaciones sobre comprensión del contraste de hipótesis. Ello es debido a que las tablas de contingencia, además de servir para estudiar la asociación entre las variables, pueden usarse para trabajar probabilidades marginales, conjuntas o condicionales. Además el análisis de la asociación en dichas tablas, a nivel inferencial requiere la comprensión del contraste Chi-cuadrado.

Respecto al primer tipo de estudios, la principal conclusión es que las tablas de contingencia han sido poco estudiadas desde el punto de la investigación didáctica, siendo excepción los trabajos de Antonio Estepa. Pensamos que la didáctica de las tablas de contingencia es un terreno con mucho que explorar, ya que su estudio se incluye en muchas especialidades universitarias, tanto en estadística, como en otras materias.

Los resultados de estos estudios muestran poca comprensión intuitiva del problema de establecimiento de un juicio de asociación y el abordaje de dicho problema mediante estrategias incorrectas; la mayor parte de las cuáles se basan tan sólo en una parte de la información disponible (una celda o una sola distribución). Por otro lado, incluso aquellos estudiantes que llegan a utilizar todos los datos, a veces usan estrategias intuitivas que se suponen abandonadas en el periodo de operaciones formales.

Los sujetos se guían adicionalmente por teorías previas sobre el problema (en vez de usar los datos empíricos) y muestran concepciones erróneas (concepción causal, local, unidireccional y determinista) que entorpecen la realización de los juicios de asociación. Algunas de estas concepciones no mejoran con la enseñanza del tema en el

experimento de Estepa (1993), aunque dicho experimento abordó una problemática muy amplia y no se dedicaron temas específicos al estudio formal de las tablas de contingencia. Por otro lado, el estudio de enseñanza realizado por Estepa, solo aborda el tema desde el punto de vista de estadística descriptiva, sin incluir la enseñanza de los contrastes de homogeneidad o asociación en tablas de contingencia o de las diversas medidas de asociación.

Ello nos ha llevado a interesarnos por un estudio del aprendizaje de los estudiantes en una enseñanza centrada específicamente en el tema y en la que también se aborden los aspectos inferenciales. Es posible que un estudio inferencial añada dificultades no estudiadas, tanto en lo que se refiere a la comprensión general del contraste de hipótesis (Vallecillos, 1996) como otras específicas relacionadas con la comprensión de la distribución Chi-cuadrado, que no han sido analizadas desde el punto de vista didáctico. Sin embargo, podría proporcionar a los estudiantes nuevas herramientas en el análisis de las tablas de contingencia que les permitan mejorar su competencia en los juicios de asociación.

Respecto al segundo tipo de estudio, la conclusión es que el uso de las tablas de contingencia trae consigo problemas de interpretación de las probabilidades asociadas estudiado en multitud de trabajos como son los de Maury (1985), Eddy (1982), Gotzsche y Olsen (2000) o Díaz (2007). Además se complica por la dificultad de la notación y la diversidad de objetos matemáticos implícitos, que se analizaron en el Capítulo 1. Estos problemas pudieran incidir también en su comprensión de la asociación, puesto que establecer un juicio de asociación, según Inhelder y Piaget (1955) implica la comparación de dos probabilidades condicionales calculadas a partir de la tabla. Sin embargo este punto no fue tenido en cuenta en el estudio de Estepa.

Finalmente, el análisis de la investigación sobre contraste de hipótesis muestra numerosas dificultades asociadas a la comprensión, tanto de los objetos matemáticos requerido en su aplicación como de la lógica del contraste (Harradine, Batanero y Rosnam, 2011). Es sobre todo muy frecuente interpretar incorrectamente el nivel de significación y el valor p (Bakan, 1966; Birnbaum, 1982; Vallecillos, 1994; Krauss y Wassner, 2002). Pero también hay incompreensión del concepto de hipótesis y falta de discriminación entre los diferentes tipos de hipótesis (Vallecillos, 1994; Chow, 1996), así como falta de competencia en la asignación de hipótesis en un contraste (Vallecillos, 1994; Vera, Díaz y Batanero; 2011). Pero la mayor parte de investigaciones se centra en

ítems de opción múltiple y no hemos encontrado ninguna centrada en el contraste de asociación o de homogeneidad a partir de tablas de contingencia.

Todas estas reflexiones indican en interés de realizar un estudio del aprendizaje formal de la asociación con una muestra de estudiantes de Psicología (Capítulo 4) y evaluar su aprendizaje (capítulos 5 y 6). De esta evaluación se espera obtener información sobre las dificultades ligadas al estudio formal de las tablas de contingencia, que no han sido analizadas en los trabajos previos. Para comparar con el conocimiento inicial del tema por este tipo de estudiantes y además, completar el análisis de estrategias de Estepa mediante la determinación de conflictos semióticos, utilizando nuestro marco teórico, llevaremos también a cabo un estudio inicial de evaluación (Capítulo 3).

CAPÍTULO 3.

EVALUACIÓN INICIAL DE ESTRATEGIAS INTUITIVAS Y ESTIMACIÓN DE LA ASOCIACIÓN

- 3.1. Introducción
- 3.2. Objetivos
- 3.3. Cuestionario
 - 3.3.1. Selección y adaptación de ítems
 - 3.3.2. Análisis del contenido de los ítems del cuestionario
 - 3.3.3. Prueba piloto y revisión
- 3.4. Revisión del instrumento
- 3.5. Descripción de la muestra
- 3.6. Juicios de asociación
- 3.7. Estimación del coeficiente de asociación
- 3.8. Consistencia de los datos
- 3.9. Teorías previas sobre la asociación
- 3.10. Análisis de estrategias
 - 3.10.1. Niveles de elaboración de estrategias
 - 3.10.2. Metodología del análisis semiótico de estrategias
 - 3.10.3. Configuraciones cognitivas en las estrategias
 - 3.10.4. Conflictos semióticos
- 3.11. Conclusiones sobre el estudio inicial de evaluación

3.1. INTRODUCCIÓN

En este capítulo se describe un estudio inicial de evaluación, donde se analizan las respuestas de una muestra de estudiantes de psicología a un cuestionario abierto, formado por ítems adaptados de Estepa (1993). Como se indicó en el Capítulo 1, se trata de una investigación explicativa (Losada y López Feal, 2003), pues nuestra preocupación se centra en determinar los orígenes o las causas de un determinado conjunto de fenómenos (estrategias y juicios de asociación). Además es una investigación cuasi-experimental, pues la muestra es intencional, pero podemos manipular las variables independientes del cuestionario.

El principal objetivo es describir los conocimientos informales con que los estudiantes de psicología se acercan al estudio de las tablas de contingencia, al ingresar en la universidad. Puesto que se utilizan ítems de Estepa (1993) y se analizan muchas de las variables consideradas por el autor, también se compararán nuestros resultados con los suyos, tratando de ahondar en el análisis, por medio de la identificación de conflictos semióticos. Posteriormente en el Capítulo 6 se comparan los conocimientos

informales de los estudiantes mostrados en este estudio inicial con los conocimientos formales que adquieren una vez estudiado el tema.

El estudio se lleva a cabo con una muestra de estudiantes de tres universidades y tamaño suficiente (414 estudiantes) para asegurar la generalizabilidad. A continuación analizamos la muestra, el cuestionario utilizado, su construcción y prueba piloto y los resultados obtenidos. Parte de dichos resultados se han publicado en Cañadas, Batanero, Contreras y Arteaga (2011, 2012).

3.2. OBJETIVOS

Como se indicó en el Capítulo 1, uno de nuestros objetivos era realizar un estudio de evaluación sobre las estrategias intuitivas, juicios de asociación y estimación de la intensidad de la asociación en tablas contingencia con una muestra reducida de ítems y estudiantes, antes de haber estudiado formalmente el tema. Más precisamente, el tercer objetivo de nuestro estudio fue el siguiente:

- *O3: Realizar un estudio de evaluación sobre las estrategias intuitivas, juicios de asociación y estimación de la intensidad de la asociación en tablas contingencia en los estudiantes de psicología.*

En este Capítulo se trata de cumplir dicho objetivo, que se descompondrá en los siguientes objetivos parciales:

O3.1: Construir o adaptar un instrumento de evaluación que permita, en un tiempo razonable, acceder a los conocimientos intuitivos de los estudiantes antes de la enseñanza y tenga en cuenta las principales variables de tarea que afectan a la interpretación de las tablas de contingencia. El instrumento es necesario para realizar el estudio de evaluación. Se partirá del cuestionario de Estepa (1993) y Estepa y Batanero (1995), seleccionando y adaptando los ítems adecuados a nuestro propósito.

O3.2: Realizar un estudio cuantitativo con dicho instrumento de los juicios de asociación, estimación de su intensidad y explicitación de teorías previas en una muestra adecuada de estudiantes. Para este estudio se elegirá una muestra de estudiantes de psicología de tamaño suficiente y de varias universidades. Se tratará de completar el estudio anterior de Estepa (1993).

O3.3: Realizar un análisis semiótico de las estrategias usadas por los estudiantes en los juicios de asociación. Este es un nuevo tipo de análisis, basado en nuestro marco teórico, que permitirá explicar algunas de las dificultades en las estrategias y juicios de

asociación que ya fueron descritas por Estepa o por otros investigadores.

O3.4: *Comparar nuestros resultados con el estudio de Estepa (1993) con estudiantes del curso preuniversitario.*

Todos estos objetivos se abordan en el resto del capítulo.

3.3. CUESTIONARIO

El cuestionario consta de 4 ítems, 3 de tablas de contingencia 2x2 y 1 de tabla de contingencia 2x3 (Anexo 1).

3.3.1. SELECCIÓN Y ADAPTACIÓN DE ÍTEMS

Para construir el cuestionario, se comenzó seleccionando algunos de los ítems del cuestionario de Estepa (1993) y Estepa y Batanero (1995), el cual constaba de diez ítems, cinco de ellos de tablas de contingencia, en los que centramos nuestra atención. La muestra a la que pasaron su cuestionario estaba formada por 213 estudiantes del Curso de Orientación Universitaria de tres institutos de bachillerato de la provincia de Jaén. Los autores consideraron, que el estudio realizado fue una evaluación de las concepciones iniciales de los estudiantes participantes, ya que la mayoría, de los estudiantes no habían estudiado formalmente el tema de dependencia funcional y aleatoria (únicamente tres estudiantes si la habían estudiado).

De los cinco ítems del citado cuestionario que trataban de las tablas de contingencia, se seleccionaron los cuatro primeros (tres tablas 2x2 y una tabla 2x3), por considerar que eran suficientes para nuestros fines (Anexo 1). Además se llevaron a cabo una serie de modificaciones que se describen a continuación, señalando los cambios respecto a los utilizados por los autores.

Datos y contexto

En primer lugar, se realizó un estudio de los datos incluidos en las diferentes tablas de contingencia de cada uno los ítems. Puesto que, en dos de ellos, las frecuencias de las celdas eran muy pequeña y, formalmente, no permitía aplicar el contraste Chi-cuadrado, en dichos ítems se aumentó dicha frecuencia, conservando el mismo tipo de asociación y un valor aproximado del coeficiente de asociación. En el último ítem, conservando las frecuencias marginales, se modificó un poco las frecuencias dobles en las celdas, para que el signo de la asociación fuese más claro que en el estudio de Estepa (1993) y Estepa y Batanero (1995).

Por otro lado, se modificaron algunos contextos, adaptándolos a problemas del campo de psicología, para motivar la participación de los estudiantes. En caso de hacer este cambio, se procuró que la relación entre datos y teorías previas fuese la misma que la presentada en el mismo ítem en el estudio de Estepa y Batanero. Todo ello, con el fin de poder comparar los resultados de los dos estudios.

Coefficientes de asociación

Para cada uno de los ítems seleccionados se calcularon dos clases de coeficientes de asociación: asimétricos (cuando una de las variables hace el papel de independiente y la otra de dependiente) y simétricos (cuando no se diferencia entre dependiente e independiente).

Los coeficientes asimétricos calculados son, el Lambda, el coeficiente de incertidumbre, D de Sommers y Eta (Batanero y Díaz, 2008). El coeficiente Lambda mide la utilidad de la variable de la fila (o columna) en la predicción de la otra variable, puesto que proporciona el porcentaje de reducción en el error cuando las filas (o columnas) se utilizan para predecir la otra variable. El coeficiente D de Sommers mide la correlación por rangos y las variables pueden ser simétricas o dependientes. Nosotros las hemos calculado para el caso asimétrico, que tiene la misma interpretación que la anterior. El coeficiente de incertidumbre expresa el grado de incertidumbre que conseguimos reducir cuando utilizamos una variable para efectuar pronósticos sobre otra. Posee dos versiones *asimétrica* (dependiendo de cuál de las dos variables consideremos dependiente) y una *simétrica* (para cuando no hacemos distinción entre variable independiente y dependiente). Lo hemos calculado para el caso asimétrico.

El coeficiente de correlación Eta sirve para cuantificar el grado de asociación existente entre una variable cuantitativa (medida en escala de intervalo o razón) y una variable categórica (medida en escala nominal u ordinal). Se trata de un coeficiente de correlación que no supone *linealidad* y cuyo cuadrado puede interpretarse como la proporción de varianza de la variable cuantitativa que está explicada por la variable categórica. Lo utilizaremos sólo en la cuarta tabla, en que la variable “número de horas de estudio” es cuantitativa, medida en escala de razón, aunque en la tabla se presente discretizada.

Respecto a los coeficientes simétricos, calcularemos el de Contingencia, V de Cramer, Gamma condicional, R de Pearson, Tau b de Kendall y Tau c de Kendall (Ruiz-Maya, et al., 1995). Los coeficientes de contingencia y V de Cramer miden la asociación

en variables nominales. El coeficiente de contingencia se rige por las mismas reglas de la correlación y las mediciones de índice correspondiente de + 1 a - 1, pasando por el cero, donde este último significa no correlación entre las variables estudiadas y los dos primeros la correlación máxima. En esta prueba estadística existe el cero, pero no alcanza la unidad, limitación que desfavorece la prueba, pues el máximo de asociación corresponde al número de categorías de las variables. La medida V de Cramer se utiliza cuando la variable de menor nivel es nominal, con valores entre 0 y 1 en tablas 2x2 y valores entre 0 y 1 en tablas $r \times c$ con $r, c > 2$ (Aguilera, 2001).

El coeficiente R mide la correlación numérica y el resto la correlación por rangos u ordinal. Gamma es una medida de asociación simétrica entre dos variables ordinales cuyo valor siempre está comprendido entre -1 y 1. Los coeficientes Tau-b de Kendall y Tau-c de Kendall son medidas no paramétrica de la correlación para variables ordinales o de rangos, cuyo signo indica la dirección de la relación y su valor absoluto indica la magnitud. Los valores posibles van de -1 a 1, pero un valor de -1 o +1 sólo se puede obtener a partir de tablas cuadradas. El primero tiene en cuenta los empates y el segundo los ignora (Batanero y Díaz, 2008).

Tipos de covariación

Para evaluar si los estudiantes confunden asociación y causalidad, también se ha tenido en cuenta, al igual que hizo Estepa (1993) los tipos de explicaciones de la asociación considerados por Barbancho (1973). Dicho autor distingue cinco casos en el estudio de la covariación entre variables: dependencia causal unilateral, interdependencia, dependencia indirecta, concordancia y covariación causal. Para decidir en que caso nos encontramos es preciso el conocimiento del contexto del problema y de las variables.

- a. *Dependencia causal unilateral.* Este primer tipo de asociación aparece cuando cada vez que se produce una variación en la variable independiente X , también hay un cambio en la variable dependiente Y . La relación no es simétrica, pues variaciones en esta segunda variable pueden no influir en la primera. Un ejemplo es la temperatura ambiental (X), que influye en la cantidad de ropa que llevamos puesta (Y), pero aunque variamos la cantidad de ropa que llevamos puesta, no cambia la temperatura ambiental.
- b. *Interdependencia.* Cuando la influencia ente las dos variables X e Y existe, y es recíproca. Por ejemplo, el desarrollo industrial de una ciudad (X) y su nivel

económico (Y). Un cambio en cualquiera de las dos variables afecta inmediatamente a la otra.

- c. *Dependencia indirecta.* Dos variables (X e Y) pueden mostrar aparentemente una covariación entre ellas, debida a una tercera variable (Z) que es la que afecta a las dos anteriores. Así, existe una relación entre el número de teléfonos móviles (X) y el número de mujeres embarazadas (Y) de una misma ciudad. Esto se explica por una tercera variable, la población de la ciudad (Z), que al variar hace cambiar estas dos variables.
- d. *Concordancia.* Es el caso en que una serie de personas u objetos son clasificados por dos personas independientes; aunque las calificaciones son independientes, existen semejanzas, que producen una correlación. Es el caso de las puntuaciones de dos profesores de los mismos exámenes, donde no influye la nota de un profesor en la del otro.
- e. *Covariación casual.* Hay casos en que dos variables muestran una variación simultánea de la que pudiera deducirse una asociación o dependencia entre dichas variables. No obstante, tal covariación puede ser totalmente casual o accidental.

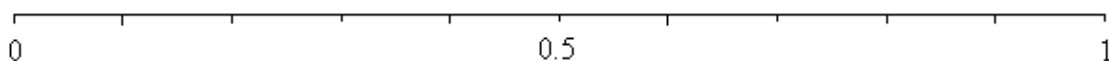
3.3.2. ANÁLISIS DEL CONTENIDO DE LOS ÍTEMS DEL CUESTIONARIO

Análisis del ítem 1

Ítem 1. Se pretende estudiar si el sufrir insomnio tiene relación con los trastornos de estrés. En una muestra de 250 personas observadas se obtuvieron los siguientes resultados:

	Padece estrés	No padece estrés
Tener insomnio	90	60
No tener insomnio	60	40

- a. Utilizando estos datos razona si el padecer o no de insomnio tiene relación con padecer estrés.
- b. Indica como has usado los datos de la tabla para llegar a tu conclusión.
- c. Asigna una puntuación entre 0 (mínimo) y 1 (máximo) según la intensidad de esta relación, marcando una cruz en el punto de esta escala que creas adecuado:



El ítem 1 es una modificación del primero de Estepa (1993) y Estepa y Batanero (1995). Se trata de un problema en el que hay independencia entre las dos variables, lo que va en contra de las teorías previas, pues en general tener estrés está asociado con el insomnio. La explicación de una posible correlación sería la dependencia causal

bilateral, pues el estrés puede ser causado por el insomnio y viceversa. Los cambios respecto a Estepa y Batanero, son ligeros, ya que utilizamos los mismos valores numéricos en la tabla, implicando independencia en ambos, y en ninguno de los dos casos hay concordancia entre los datos y las teorías previas.

Las principales diferencias son un cambio del contexto y significado de las variables, tratando de buscar un contexto que sea familiar a estudiantes de psicología. También varía el tipo de covariación, que en Estepa y Batanero era dependencia causal, siendo en nuestro caso de interdependencia. La prueba de Chi-cuadrado no es estadísticamente significativa pues se obtiene un valor de $\chi^2=0$ así como del resto de coeficientes. Al tratarse del caso de independencia total, no calculamos los coeficientes de asociación, pues todos ellos han de ser igual a cero.

Análisis del ítem 2

El ítem 2 es una modificación del segundo de Estepa (1993) y Estepa y Batanero (1995) y se trata de una situación en que hay dependencia inversa entre las dos variables, donde los datos concuerdan con las teorías previas, pues en general se considera que ser hijo único está asociado con ser un niño con más problemas que los niños que tienen hermanos. La posible correlación se explica por la dependencia causal unilateral.

Ítem 2. Un psicólogo infantil estudia si existe relación entre ser hijo único (sin hermanos) y ser un niño problemático. Para ello se han observado 250 niños obteniendo los resultados siguientes:

	Niño problemático	Niño no problemático
Tiene hermanos	40	100
No tiene hermanos	100	10

- Utilizando estos datos razona si hay alguna relación entre ser un niño problemático y tener o no hermanos.
- Indica cómo has utilizado los números de la tabla, para llegar a tu conclusión.
- Asigna una puntuación entre 0 (mínimo) y 1 (máximo) según la intensidad de esta relación, marcando una cruz en el punto de esta escala que creas adecuado:

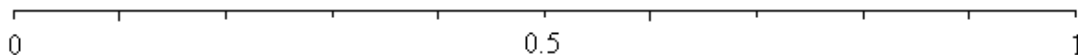


Tabla 3.3.2.1. Coeficientes de asociación asimétricos en el ítem 2

Estadístico	Con Filas Dependientes
Lambda	0,55
Coef. Incertidumbre	0,32
D de Somer	-0,62

Se han modificado los datos del enunciado original para obtener un mayor valor de coeficiente de correlación y también mayor tamaño de muestra. Asimismo se cambió el contexto original del ítem, para hacerlo más familiar a los estudiantes de psicología. La prueba de Chi-cuadrado fue estadísticamente significativa pues se obtiene un valor de $Chi=97,15$, con un grado de libertad y $p=0,0000$. Por tanto, podemos rechazar la hipótesis de que las filas (tener hermanos o ser hijo único) y columnas (niños con o sin problemas) son independientes con un nivel de confianza superior a 99,99%.

En las Tablas 3.3.2.1 y 3.3.2.2 se incluyen algunos de los coeficientes de contingencia, calculados para este ítem. En la Tabla 3.3.2.1 se presentan los coeficientes asimétricos, que varían bien entre -1 y 1 cuando tienen signo. En nuestro caso, la variable independiente sería la variable en filas. Los valores oscilan entre 0,32 y 0,62, y en caso de tener signo es negativo, pues la presencia de un carácter (tener hermanos) implica menor frecuencia del otro (ser problemático).

Tabla 3.3.2.2. Coeficientes de asociación simétricos en el ítem 2

Estadístico	Valor	P-Valor
Coef. Contingencia	0,53	
V de Cramer	0,62	
Gamma condicional	-0,92	
R de Pearson	-0,62	0,00
Tau b de Kendall	-0,62	0,00
Tau c de Kendall	-0,61	

En la Tabla 3.3.2.2 se incluyen los coeficientes de asociación simétricos (es decir cuando no se considera diferencia entre variable dependiente e independiente). Cuando los p-valores son inferiores a 0,05 indican una asociación significativa entre filas y columnas con un nivel de confianza del 95%. Observamos que los valores oscilan entre 0,53 (coeficiente de contingencia) y -0,92 (Gamma condicional) y en caso de tener signo son negativos. En resumen, la asociación entre las variables es inversa y moderada-alta.

Análisis del ítem 3

El ítem 3 es una modificación del tercero de Estepa (1993) y Estepa y Batanero (1995) y presenta un problema en el que no hay influencia de teorías previas sobre la asociación entre las variables. La explicación de una posible correlación sería una tercera variable en juego. Se han modificado los datos multiplicando todas las celdas por diez, siendo la correlación de la misma intensidad, pero estadísticamente más

significativa, al haber más casos en todas las celdas. La prueba de Chi-cuadrado fue estadísticamente significativa, pues se obtiene un valor de $Chi=133,93$, con un grado de libertad y $p=0,0000$. Por ello podemos rechazar la hipótesis de que las filas (forma de vida sedentaria o no) y columnas (ingerir una dieta blanda) son independientes con un nivel de confianza del 99%.

Ítem 3. Se quiere saber si sufrir o no de alergia tiene relación con llevar una vida sedentaria (llevar una vida sin realizar ningún tipo de ejercicio físico). Para ello obtuvimos los siguientes resultados de 300 sujetos:

	Sufre alergia	No sufre alergia
Forma de vida sedentaria	130	30
Forma de vida no sedentaria	20	120

- Utilizando estos datos razona si padecer alergia tiene o no relación con llevar una vida sedentaria en estos sujetos.
- Indica como has usado los números de la tabla, para obtener tu conclusión.
- Asigna una puntuación entre 0 (mínimo) y 1 (máximo) según la intensidad de esta relación, marcando una cruz en el punto de esta escala que creas adecuado:

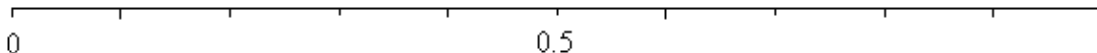


Tabla 3.3.2.3. Coeficientes de asociación asimétricos en el ítem 3

Estadístico	Con Filas Dependientes
Lambda	0,64
Coef. Incertidumbre	0,35
D de Somer	0,67

Los coeficientes asimétricos (Tabla 3.3.2.3) son moderados o moderados-altos y de signo positivo, oscilando entre 0,35 y 0,67. En la Tabla 3.3.2.4 se incluyen los coeficientes de asociación simétricos, cuyos valores oscilan entre 0,56 (coeficiente de contingencia) y 0,93 (Gamma condicional) y tienen signo positivo. En resumen, la asociación entre las variables es directa y moderada-alta.

Tabla 3.3.2.4. Coeficientes de asociación simétricos en el ítem 3

Estadístico	Valor	P-Valor
Coef. Contingencia	0,56	
V de Cramer	0,67	
Gamma condicional	0,93	
R de Pearson	0,67	0,00
Tau b de Kendall	0,67	0,00
Tau c de Kendall	0,67	

Análisis del ítem 4

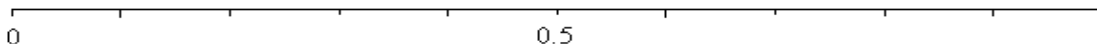
El ítem cuarto es el mismo que el de Estepa (1993) y Estepa y Batanero (1995). Se trata de un ítem en el que hay dependencia directa entre las dos variables, lo que va a

favor de las teorías previas, pues en general se considera que al incrementar el tiempo dedicado a preparar el examen está asociado con aprobar. La explicación de una posible correlación sería la dependencia causal unilateral, pues el aprobar un examen puede ser por dedicar más tiempo de preparación del examen.

Ítem 4. La siguiente tabla nos indica el número de estudiantes que aprobaron o suspendieron un examen teniendo en cuenta el tiempo que cada estudiante dedicó a prepararlo

	Menos de 5 h	Entre 5 y 10 h	Más de 10 h	TOTAL
Aprobados	5	15	51	71
Suspensos	20	7	2	29
Total	25	22	53	100

- Utilizando estos datos razona si aprobar o suspender el examen tiene relación con el tiempo dedicado a prepararlo.
- Indica como has usado los números de la tabla, para llegar a tu conclusión.
- Asigna una puntuación entre 0 (mínimo) y 1 (máximo) según la intensidad de esta relación, marcando una cruz en el punto de esta escala que creas adecuado:



No se han realizado cambios significativos en el ítem. La prueba de Chi-cuadrado fue estadísticamente significativa ($Chi=48,05$, con dos grados de libertad y $p=0,0000$). Podemos, por tanto, rechazar la hipótesis de que las filas (aprobar o suspender) y columnas (horas empleadas en la preparación del examen) son independientes con un nivel de confianza del 99%. En la Tabla 3.3.2.5 se incluyen algunos de los coeficientes de contingencia, que muestran valor menos intenso que los ítems anteriores, variando desde 0,35 a 0,52 y siendo positivos.

Tabla 3.3.2.5. Coeficientes de asociación asimétricos en el ítem 4

Estadístico	Con Filas Dependientes
Lambda	0,52
Coef. Incertidumbre	0,42
D de Somer	0,35
Estadístico Eta	0,38

Tabla 3.3.2.6. Coeficientes de asociación simétricos en el ítem 4

Estadístico	Valor	P-Valor
Coef. Contingencia	0,57	
V de Cramer	0,69	
Gamma condicional	0,61	
R de Pearson	0,38	0,00
Tau b de Kendall	0,43	0,00
Tau c de Kendall	0,43	

En la Tabla 3.3.2.6 se incluyen los coeficientes de asociación simétricos, que

oscilan entre 0,38 (*R* de Pearson) y 0,69 (*V* de Cramer) y son positivos. En resumen, la asociación entre las variables es directa y menos intensa que las anteriores.

VARIABLES DE TAREA EN LOS ÍTEMS

En resumen, las variables de tarea (Kilpatrick, 1978) tenidas en cuenta en el momento de la elaboración del cuestionario han sido las siguientes:

1. *Tipo de dependencia estadística entre las variables*, incluyéndose los tres casos posibles en nuestros ítems, es decir, dependencia directa, dependencia inversa e independencia en las tablas 2x2.
2. *Intensidad de la dependencia*, medida mediante el coeficiente Phi, que coincide con el de Pearson en tablas 2x2 y con el coeficiente *V* de Cramer en tablas 2x3. la intensidad de la asociación varía de 0 (ítem 1) a 0,67 (ítem 3), teniendo un ítem de intensidad moderada-baja y dos de intensidad moderada-alta.
3. *Concordancia entre los datos y las teorías previas sugeridas por el contexto del problema*. Se han elegido ítems en los que coinciden los resultados con las teorías previas, que no coinciden y otras donde no hay teorías previas.
4. *Tipo de covariación*, utilizando algunas de las categorías descritas por Barbancho (1973), específicamente las siguientes: dependencia causal unilateral, cuando una variable puede considerarse causa de la otra; interdependencia, cuando hay una relación de dependencia mutua y dependencia indirecta, si las dos variables dependen de una tercera que es la que ocasiona la covariación.

En la Tabla 3.3.2.7, se presenta un resumen de los valores de las *variables de tarea* empleadas en todos nuestros ítems

Tabla 3.3.2.7. Variables de tarea

	Tabla 2x2		Tabla 2 x3	
	Item 1	Item 2	Item 3	Item 4
Dependencia	Independencia	Inversa	Directa	Directa
Coefficiente correlación	0	-0,62	0,67	0,37
Concuerda con t. previa	No	Si	No hay teoría	Si
Tipo covariación	Interdependencia	Causal unilateral	Tercera variable	Causal unilateral

3.3.3. PRUEBA PILOTO Y REVISIÓN

Finalizada la primera versión del cuestionario, se realizó una prueba piloto con algunos estudiantes de Psicología, dentro de la asignatura de “Análisis de datos”, con objeto de obtener información empírica sobre sus características y ver si resultaría de

utilidad en cuanto a si nos aportaba la información que requerimos para los objetivos pretendidos. Las características de los estudiantes fueron en todo similares a las de los que respondieron el cuestionario definitivo (edad, curso, estudios previos, distribución por género).

Para ello se pidió a los estudiantes, que participaron voluntariamente, responder al cuestionario con el mayor detalle posible e indicar también los puntos de la redacción que les pareciesen confusos. De este modo quisimos comprobar la legibilidad de los ítems y el tiempo que necesitarían para completar el cuestionario, en caso de ser pertinente la investigación. Al mismo tiempo se deseaba analizar sus limitaciones, para identificar aquellos puntos en que sería necesario mejorarlo.

Recogidos los datos se analizaron los juicios de asociación, estimación del coeficiente y estrategias utilizadas por estos estudiantes, obteniendo los resultados que se describen a continuación (publicados en Cañadas, 2011):

- Los estudiantes mostraron interés por el tema, indicando que no habían realizado anteriormente ejercicios similares; todos respondieron a los diferentes apartados de los cuatro ítems.
- Los estudiantes completaron el cuestionario en un tiempo aproximado de una hora, por lo cual se decidió que el tiempo final que se utilizaría en la muestra definitiva sería de hora y media, para evitar que algún estudiante no pudiese completarlo por falta de tiempo.
- Se hicieron sugerencias de mejora de la redacción del enunciado de uno de los ítems y de la forma en que se planteaba el apartado c) en todos ellos.
- Se analizaron los juicios de asociación dados por los estudiantes. Se observó que se considera mayoritariamente la existencia de asociación en todos los ítems, con mayor frecuencia en los casos en que existe relación en los datos, pero también en el ítem 1, que corresponde independencia perfecta. Pensamos por tanto, que el ítem 1 era útil para detectar la correlación ilusoria.
- Respecto a la precisión de la estimación del coeficiente, es también el ítem 1 en el que el valor medio estimado se aleja más del coeficiente de asociación, posiblemente por las teorías previas y además tener la mayor frecuencia en la celda a de la tabla. En el caso de dependencia directa (ítem 3), los estudiantes se aproximan casi perfectamente al coeficiente de asociación. Para el ítem 2 de dependencia inversa, se ve un acercamiento de los estudiantes al coeficiente, aunque no tan bueno como en el otro caso. En general, comparando los ítems donde hay

teorías previas, vemos que los resultados sobrestiman el valor del coeficiente, siempre que hay una teoría sobre la existencia de asociación (ítem 1, 2 y 4) mientras dan una estimación casi perfecta cuando no tienen teorías.

- Se analizaron también la corrección de las estrategias empleadas, observando tanto estrategias correctas y parcialmente correctas como incorrectas, por ejemplo, el uso de la celda de mayor frecuencia, de una sola distribución condicional o comparar frecuencias dobles entre sí.

En consecuencia, deducimos que los ítems eran adecuados para obtener la información que necesitábamos, por lo que se decidió usar únicamente estos cuatro ítems en el cuestionario, con ligeras modificaciones que se describen a continuación.

3.4. REVISIÓN DEL INSTRUMENTO

Para mejorar el cuestionario, teniendo en cuenta lo observado en el estudio piloto, así como las sugerencias de los estudiantes, se realizaron algunos cambios superficiales, que son los siguientes:

- El primer cambio, se sitúa en los datos del estudiante, donde añadimos una pregunta pidiendo su nombre y apellido, para motivarlos a responder (pues se les puntuaba en la asignatura su colaboración).
- Se incluyó un apartado para pedir a los estudiantes que, junto a la nota de acceso de selectividad, indicaran la fecha de examen. Esto fue necesario ya que la escala de valoración fue diferente en distintas convocatorias (desde diez hasta catorce), y de esta forma se pudo poner la nota de acceso en la misma escala.
- Se mejoró la redacción de los enunciados, en los casos que fue necesario para mejorar la comprensión, quedando el enunciado final el que hemos presentado en este capítulo.
- Por último, en el apartado (c) de todos los ejercicios, se añadió el significado exacto de marcar el valor 0, y el valor 1 en la escala. Por ejemplo, para el ítem 1:

“0: Indica que no existe ninguna relación entre tener insomnio y padecer estrés

1: Indica que, sabiendo si una persona padece o no insomnio, podemos decir con seguridad si tiene o no estrés”.

3.5. DESCRIPCIÓN DE LA MUESTRA

La muestra estuvo formada por un total de 414 estudiantes de primer año de la Licenciatura en Psicología¹ (62 de la Universidad de Huelva, 237 de la Universidad de Granada y 115 de la Universidad de Almería), que cursaban una asignatura de “Análisis de Datos I”, cuyos profesores colaboraron con nosotros. También colaboraron en la recogida de datos los estudiantes que participaron en la prueba piloto.

Esta asignatura se enmarcaba dentro del área de Métodos de Investigación en las Ciencias del Comportamiento, que incluía en la fecha en que se tomaron los datos varias asignaturas troncales y obligatorias de universidad, al igual que en los nuevos planes, aunque las asignaturas han variado un poco. Tanto en el plan anterior, como en el actual, se ofrece a los estudiantes un cuerpo de conocimientos teóricos y prácticos cuya comprensión y utilización les permitirá analizar y entender los resultados obtenidos, tanto en la investigación como en la práctica psicológica, así como generar sus propios análisis de datos en aquellos campos que resulten de su interés.

Tabla 3.5.1. Frecuencia y porcentaje del sexo de los estudiantes de la muestra

Sexo	Frecuencia	Porcentaje
Hombre	96	23,19
Mujer	313	75,6
No contesta	5	1,21
Total	414	100

Los objetivos y competencias esperados son muy similares a los de la actual materia: *Descripción y exploración de datos en Psicología*, que se describe con detalle en el capítulo 4, por lo cuál no los repetimos acá. Los contenidos de la asignatura incluyen la estadística descriptiva univariante, teorema de Bayes y estudio de las distribuciones de probabilidad normal, *t* de Student, Chi-cuadrado y *F* de Snedecor.

Los estudiantes no habían estudiado todavía el tema de las tablas de contingencia, por lo que los conocimientos que pueden utilizar son los adquiridos en Bachillerato o bien estrategias intuitivas. En nuestra muestra de 414 estudiantes, aparecen 96 hombres y 313 mujeres (5 estudiantes no responden), es decir en la muestra el 23,19% estaba constituido por hombres, el 75,6% por mujeres y 1,21% no responden (ver Tabla 3.5.1). Podemos destacar que en la variable sexo hay casi el triple de mujeres que de hombres, que es la proporción normal en Psicología

También consideramos interesante observar otras variables como la especialidad

¹ Que se transforma en Grado de Psicología el año siguiente al que se realizó la toma de datos.

de bachillerato de donde precedían los estudiantes (ver Tabla 3.5.2), si el estudiante estaba realizando la asignatura por segundo año (ver Tabla 3.5.3) o la nota de acceso a la universidad con la que habían entrado en psicología (Figura 3.5.1).

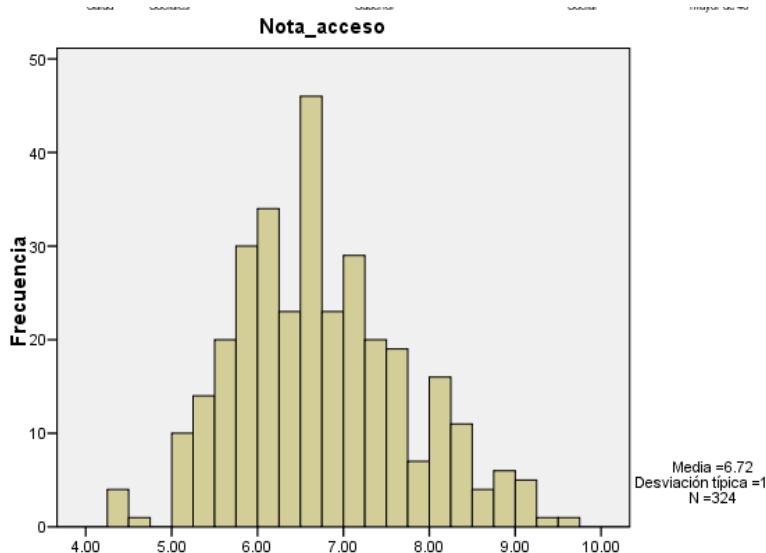
Tabla 3.5.2. Frecuencia y porcentaje de la especialidad de Bachillerato en los estudiantes de la muestra

	Frecuencia	Porcentaje
Ciencias de la Salud	127	30,7
Ciencias Sociales	217	52,5
Tecnológico	27	6,5
Internacional	4	1
Ciclo Superior	9	2,2
Erasmus	1	0,2
Arte	7	1,7
Integración Social	2	0,5
Diplomatura	1	0,2
Acceso mayor 40 años	1	0,2
No contesta	18	4,3
Total	414	100

Tabla 3.5.3. Frecuencia y porcentaje de los estudiantes de la muestra que repiten curso

	Frecuencia	Porcentaje
Si	22	5,3
No	374	90,4
No contesta	18	4,3
Total	414	100

Figura 3.5.1. Nota de acceso de los alumnos participantes



Los estudiantes entran en Psicología principalmente con notas bajas, ya que se ve la disminución de la frecuencia cuanto mayor es la nota de acceso; esto es debido a que no hay nota de corte. Estos estudiantes proceden principalmente del Bachillerato de Ciencias Sociales y algo menos del Bachillerato de Ciencias de la Salud, siendo casi

nula la presencia de estudiantes procedentes del Bachillerato Tecnológico. Y finalmente, aproximadamente una sexta parte de los participantes son repetidores. En consecuencia, es de esperar que los conocimientos estadísticos no sean muy fuertes.

3.6. JUICIOS DE ASOCIACIÓN

En el apartado a) de cada ítem se pide a los estudiantes un juicio sobre la posible relación entre las variables presentadas en la tabla. Para evaluar la competencia de los estudiantes en este punto, se presenta en la Tabla 3.6.1 el porcentaje de estudiantes que considera (o no) la existencia de relación entre las variables. Como se observa en dicha tabla, la mayoría de los estudiantes indica la existencia de asociación en todos los ítems, con mayor frecuencia en los casos en que existe relación en los datos, pero también en el ítem 1, donde los datos presentan independencia perfecta. En lo que sigue analizamos los resultados en cada ítem, comparando con el estudio de Estepa (1993).

Tabla 3.6.1. Frecuencia (y porcentaje) de juicios sobre la asociación entre las variables

Ítem	Consideran relación	Consideran independencia	No contestan	Coefficiente de asociación	Teorías previas
1	323 (78,1)	90 (21,7)	1 (0,2)	0	En contra
2	398 (96,1)	14 (3,4)	2 (0,5)	-0,62	A favor
3	386 (93,2)	24 (5,8)	4 (1)	0,67	No hay
4	402 (97,1)	4 (1)	8 (1,9)	0,37	A favor

Ítem 1: Independencia en tablas 2x2, con teorías en contra. A pesar de la independencia de los datos, una abrumadora mayoría cree que hay asociación. Este resultado puede explicarse por el mecanismo de correlación ilusoria, descrito por Chapman y Chapman (1969), Jennings, Amabile y Ross (1982), Alloy y Tabacnick (1984), Wright y Murphy (1984) y Murphy y Medin (1985), pues los estudiantes conocen que el estrés puede provocar pérdida de sueño; es un contexto muy conocido para ellos.

Al comparar con el ítem equivalente en el estudio de Estepa, el 39,4% de estudiantes consideró la independencia, siendo menor el porcentaje en nuestro caso (21,7%). Como se indicó al analizar este ítem, la única diferencia con el de Estepa es el contexto, más adecuado al campo de la psicología en nuestro caso. Nuestros estudiantes muestran un mayor efecto de las teorías previas en este ítem, así como la concepción causal de la asociación (Estepa, 1003), ligando los conceptos de asociación y causalidad. Los estudiantes manifestaron estas teorías previas y razonamientos causales en sus respuestas, como se verá posteriormente en las secciones 3.9 y 3.10.

También en las dos investigaciones aparece la frecuencia mayor en la celda *a*, lo cual puede ser una explicación del bajo número de aciertos, ya que en muchas investigaciones se han encontrado sujetos que únicamente se fijan en la celda donde se presentan los dos caracteres (Smedlund, 1963; Beyth y Marom, 1982; Shaklee y Mins, 1982, Yates y Curley, 1986). Además, para resolver el ítem, es preciso un correcto empleo del razonamiento proporcional ya que se puede ver como un problema de comparación de dos probabilidades.

Ítem 2: Asociación Inversa en tablas 2x2. La proporción de estudiantes que indica que hay relación entre las variables fue muy alta en nuestro caso (96,1% de los estudiantes). En el ítem equivalente a este en el estudio de Estepa (1993) el autor preguntó por el tipo de asociación (directa o inversa); el 50,7% de los estudiantes fueron capaces de reconocer la dependencia inversa, el 11,7% de los estudiantes consideran que existe una asociación directa y el resto no contesta o considera independencia.

Los resultados son entonces comparativamente mejores en nuestro estudio, donde hemos aumentado la intensidad y colocado en cada celda valores superiores a diez, para asegurar la aplicabilidad del estadístico Chi-cuadrado. Todos estos cambios han llevado a que apenas se haya presentado en nuestros estudiantes la concepción unidireccional de la asociación descrita por Estepa (1993). En nuestro caso no preguntamos el signo de la asociación, y el porcentaje de estudiantes que considera asociación es mucho mayor. Esto puede explicar las diferencias, así como la familiaridad del contexto, el hecho de que las teorías previas apoyan los datos y la mayor intensidad de la asociación.

Ítem 3: Asociación directa en tablas 2x2. Este ítem corresponde a dependencia directa, con teorías previas a favor de los datos, por lo cual los resultados son muy buenos, bastante similares al estudio de Estepa (91,1% de estudiantes consideran asociación directa). La diferencia con nuestro ítem se debe a que hemos cambiado el contexto pues Estepa usaba un contexto médico (trastornos digestivos e ingerir una dieta blanda). También se ha multiplicado todas las celdas por diez, lo que puede ser la razón de un mayor porcentaje de aciertos.

Ítem 4: Asociación directa en tablas rxc. En este ítem (aprobados y tiempo dedicado a preparar un examen) y teorías previas a favor de los datos, los estudiantes aciertan en su mayoría considerando asociación, también con mejores resultados que en el estudio de Estepa (87,3% de estudiantes considera la asociación). La diferencia en nuestro ítem en la intensidad de la dependencia, ya que modificamos un poco los datos,

siendo en su investigación la intensidad de 0,43 y en nuestro caso de 0,37. El contexto del problema fue el mismo en ambas investigaciones.

Comparación por universidades

Con objeto de comprobar la homogeneidad de las tres muestras, respecto a los juicios de asociación en los ítems, se descomponen los resultados según universidad participante. En la Tabla 3.6.2 se presenta el porcentaje de estudiantes que considera asociación, según universidad, observando que los resultados son muy similares en todas ellas, a pesar de que la enseñanza de estadística en cada grupo, presumiblemente haya tenido variación. Parece en consecuencia que los resultados podrían ser generalizables a otras universidades; además, la homogeneidad de las muestras nos permitiría en el resto del estudio tratarlas como una sola.

Tabla 3.6.2. Frecuencia (y porcentaje) de estudiantes que considera asociación por ítem y muestra

Ítem	Almería (n=115)	Granada (n=237)	Huelva (n=62)	Total (n=414)	Coefficiente de asociación	Teorías previas
1	87 (75,7)	194 (81,9)	42 (67,7)	323 (78,1)	0	En contra
2	108 (93,9)	230 (97,1)	60 (96,8)	398 (96,1)	-0,62	A favor
3	107 (93,1)	226 (95,4)	53 (85,5)	386 (93,2)	0,67	No hay
4	110 (95,7)	235 (99,2)	57 (91,9)	402 (97,1)	0,37	A favor

Para comprobar si las pequeñas diferencias en las proporciones de respuestas a cada ítem son o no estadísticamente significativas se realizó un contraste de homogeneidad de muestras Chi-cuadrado obteniendo un valor $Chi= 0,99$ con 6 g.l, lo que corresponde a un valor $p=0,9861$. Por tanto, los resultados no son estadísticamente significativos y más aún el valor p tan cercano a la unidad indica que las muestras son homogéneas en cuanto a los juicios de asociación, de lo que se deduce que los resultados no dependen de la Universidad de procedencia del estudiante.

3.7. ESTIMACIÓN DEL COEFICIENTE DE ASOCIACIÓN

En el apartado (c) de cada ítem se pedía al estudiante una puntuación entre 0 y 1, según la intensidad que ellos percibiesen en la asociación entre las variables. La exactitud en la estimación de este coeficiente sólo fue analizada por Estepa (1993) en una muestra piloto de 50 estudiantes, pero no en su estudio definitivo. Tampoco analizó la influencia de la universidad de procedencia, juicio de asociación o estrategia seguida sobre dichos juicios. Nosotros realizaremos estos tipos de análisis en lo que sigue.

Para ello, hemos codificado mediante un valor numérico la intensidad estimada

por los estudiantes. En algunos casos, los estudiantes han indicado el valor numérico; en el resto, hemos transformado el punto dibujado en la escala proporcionada a un valor numérico, midiendo con una regla milimetrada la distancia al origen y transformando la escala convenientemente. En la Tabla 3.7.1 presentamos el valor medio obtenido en el conjunto de la muestra, y en cada una de las universidades. Dicho valor sería una estimación del coeficiente de asociación (no se solicitó especificar el signo).

La mayor precisión de la estimación se da el ítem 3, que corresponde a la dependencia directa y con un valor alto del coeficiente de asociación, donde, además los estudiantes no tienen teoría al respecto. Observamos que el valor medio estimado es prácticamente idéntico al empírico. Hay una sobrestimación del coeficiente en los otros tres ítems para sendas provincias y la muestra global, observándose el efecto de las teorías previas de los estudiantes.

Tabla 3.7.1. Estimación del coeficiente de asociación por ítem y muestra

Ítem	Media estimada				Coeficiente de asociación	Teorías previas
	Almería (n=115)	Granada (n=237)	Huelva (n=62)	Total (n=414)		
1	0,51	0,47	0,44	0,47	0	En contra
2	0,78	0,72	0,73	0,73	-0,62	A favor
3	0,75	0,68	0,68	0,70	0,67	No hay
4	0,84	0,81	0,81	0,82	0,37	A favor

Como podemos observar en la tabla 3.7.1, los estudiantes proporcionan siempre un coeficiente de asociación alto o moderado, por el fenómeno, ya descrito, de correlación ilusoria. En el ítem 1, que corresponde al caso de independencia perfecta, el valor medio estimado osciló entre 0,44 (Huelva) y 0,51 (Almería), obteniéndose un valor medio global estimado de 0,47, aunque, de hecho, al haber independencia en los datos el verdadero valor del coeficiente es 0. Por otro lado, en este ítem la mayor frecuencia aparece en la celda *a*, es decir la primera de las frecuencias absolutas de la tabla que corresponde a la presencia de los dos caracteres y que, según varios autores (Smedlund, 1963; Beyth y Marom, 1982; Shaklee y Mins, 1982, Yates y Curley, 1986) tiene mayor impacto sobre la atención de los sujetos que las otras celdas. Todo ello podría haber influido en la estimación de una asociación no existente en los datos.

Para el ítem 2 correspondiente a una dependencia inversa, cuyo coeficiente de asociación tiene un valor absoluto próximo al del ítem 3 (cuya dependencia es directa), se estima un coeficiente de asociación próximo al empírico en todas las universidades, algo inferior en Granada. De este modo, en nuestros estudiantes no aparece un

porcentaje apreciable de concepción unidireccional descrita por Estepa (1993), ni en cada muestra parcial, ni en el global, pues además, la mayoría de los estudiantes indica que hay asociación (no se pidió el signo de la asociación). Este mejor resultado puede deberse a que las teorías previas estaban a favor de la asociación. Nuestros resultados contradicen el estudio de Erlick y Mills (1967) quien indica que la asociación negativa se estima como muy próxima a cero.

En el caso de dependencia directa (ítem 3), los estudiantes se aproximan casi perfectamente al coeficiente de asociación. En este ítem no hubo teorías previas sobre el tema y los estudiantes se guiaron más por la observación de los datos, que muestran con facilidad la dependencia, ya que seleccionamos un coeficiente de asociación alto. Estos resultados coinciden con los de Jennings, Amabile y Ross (1982), Wrigth y Murphy (1984), Alloy y Tabacnick (1984) y Meiser y Hewstone (2006), quienes indican que la estimación de la asociación es más precisa si las personas no tienen teorías respecto al tipo de asociación sobre los datos.

En el ítem 4, correspondiendo a una tabla 2x3 con una asociación positiva, se puede notar la gran diferencia que hay entre la estimación y su valor verdadero. Los estudiantes sobrestiman la asociación, siguiendo sus teorías previas, que coinciden con el tipo de asociación en los datos. Estas teorías se manifestaron en los estudiantes al ser un contexto tan cercano a ellos (horas de estudio-aprobar o suspender), y como consecuencia se han dejado llevar por su experiencia personal. Esto ocurre en todas las universidades.

En general, comparando los ítems donde hay teorías previas, vemos que los resultados sobrestiman el valor del coeficiente, siempre que hay una teoría sobre la existencia de asociación (ítem 1, 2 y 4), mientras dan una estimación casi perfecta cuando no tienen teorías, de acuerdo con lo que indican Chapman y Chapman (1969).

Los estudiantes de Granada y Huelva en todos los ítems estiman en promedio un coeficiente más bajo que los estudiantes de Almería. Para analizar si la diferencia es estadísticamente significativa se realizó un Análisis de varianza (Tabla 3.7.3). Observamos que, aunque el resultado es estadísticamente significativo, la significación es pequeña, apenas un poco menor al 5%.

Tabla 3.7.3. Resultados del Manova

	Valor	F	Gl de la hipótesis	Gl del error	Significación
Lambda de Wilks	,960	2,017(a)	8,000	780,000	,042

Tabla 3.7.4. Pruebas de los efectos inter-sujetos

Fuente	Variable dependiente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo corregido	I1_coeficiente	,240(a)	2	,120	2,068	,128
	I2_coeficiente	,319(b)	2	,160	4,248	,015
	I3_coeficiente	,375(c)	2	,188	3,819	,023
	I4_coeficiente	,035(d)	2	,018	,733	,481
Intersección	I1_coeficiente	62,361	1	62,361	1073,916	,000
	I2_coeficiente	158,329	1	158,329	4214,912	,000
	I3_coeficiente	142,356	1	142,356	2896,487	,000
	I4_coeficiente	188,974	1	188,974	7805,434	,000
Provincia	I1_coeficiente	,240	2	,120	2,068	,128
	I2_coeficiente	,319	2	,160	4,248	,015
	I3_coeficiente	,375	2	,188	3,819	,023
	I4_coeficiente	,035	2	,018	,733	,481
Error	I1_coeficiente	22,821	393	,058		
	I2_coeficiente	14,763	393	,038		
	I3_coeficiente	19,315	393	,049		
	I4_coeficiente	9,515	393	,024		
Total	I1_coeficiente	111,944	396			
	I2_coeficiente	231,739	396			
	I3_coeficiente	216,442	396			
	I4_coeficiente	274,465	396			

a R cuadrado = ,010 (R cuadrado corregida = ,005)

b R cuadrado = ,021 (R cuadrado corregida = ,016)

c R cuadrado = ,019 (R cuadrado corregida = ,014)

d R cuadrado = ,004 (R cuadrado corregida = -,001)

Al analizar las pruebas de efecto inter-sujeto (Tabla 3.7.3), se observa que sólo son significativas las diferencias en los ítems 2 y 3 y, aún así, el efecto es pequeño pues el coeficiente de correlación al cuadrado toma un valor apenas mayor del 1%. No observamos, por tanto, significación práctica en las diferencias de los ítems, por lo que se decide no realizar nuevos análisis estadísticos y aceptar las tres muestras como equivalentes, respecto a la estimación del coeficiente y en el resto del estudio.

3.8. CONSISTENCIA DE LOS DATOS

Analizados separadamente el juicio de asociación y la estimación del coeficiente, en este apartado estudiamos su consistencia, así como la existente entre el tipo de estrategia (correcta, parcialmente incorrecta o incorrecta) utilizada por el estudiante y la precisión en la estimación del coeficiente de asociación.

Consistencia entre juicio de asociación y estimación del coeficiente

En primer lugar, nos preguntamos si la estimación que realizan los estudiantes del coeficiente de asociación es consistente con su percepción subjetiva de la asociación, que viene dada por su juicio de asociación positivo o negativo en el ítem. Para estudiar esta consistencia, en la Tabla 3.8.1 se presentan los valores medios y error típico del coeficiente estimado en cada ítem, según si el juicio de asociación haya sido positivo (se indica que hay asociación en los datos) o negativo (se niega la asociación). De ser los datos consistentes, debiéramos esperar mejor precisión, cuando el juicio es correcto.

Asimismo se presentan las diferencias de valores medios obtenidos para el coeficiente en estos dos grupos de estudiantes, junto con el error típico de la diferencia de medias en los dos grupos, el valor t en el contraste de diferencias de medias en muestras independientes y el valor p correspondiente a dicho contraste.

Tabla 3.8.1. Estimación de la asociación según consideran o no asociación en los datos

Ítem	Consideran asociación				Diferencia Medias	E. típico diferencia	t	p
	Si		No					
	Media	E. Típico	Media	E. Típico				
1	0,56	0,01	0,18	0,02	0,38	0,02	17,00	0,000
2	0,75	0,01	0,31	0,55	0,44	0,05	8,76	0,000
3	0,73	0,01	0,22	0,04	0,51	0,04	12,10	0,000
4	0,82	0,01	0,31	0,08	0,51	0,07	6,89	0,000

Observamos una diferencia estadísticamente muy significativa en todos los ítems, y un valor medio siempre superior en el grupo de estudiantes que da un juicio positivo de asociación, que en el grupo de los que dan un juicio negativo. Ello indica que el valor medio estimado para el coeficiente de asociación es mayor en el grupo de estudiantes que da un juicio positivo de asociación, como era nuestra hipótesis.

Consecuentemente, hay una alta consistencia entre el juicio de asociación y el valor estimado del coeficiente que es siempre más alto (correspondiendo a valores de correlación alta o moderada) en los estudiantes que perciben asociación en los datos. Por el contrario los valores medios de los coeficientes corresponden a asociaciones bajas en los que dan un juicio negativo.

Consistencia entre corrección de la estrategia y estimación del coeficiente

Por otro lado, se ha analizado la existencia de diferencias en la estimación de la asociación, según el tipo de estrategia que emplean los estudiantes (incorrecta,

parcialmente correcta ó correcta), pues habría que esperar que a mayor corrección de la estrategia fuera mejor la estimación del coeficiente. Aunque posteriormente haremos un estudio más detallado de estrategias, en este punto sólo tenemos en cuenta su corrección. Son estrategias correctas aquellas que utilizan todos los datos de la tabla mediante la comparación de probabilidades. Por ejemplo, en el ítem 4 un estudiante indica:

“Para menos de 5 horas: el 20% aprueba y el 80% suspende; Entre 5 y 10 horas: el 66% aprueba y el 33% suspende; Para más de 10 horas: el 96% aprueba y el 3% suspende. En la gráfica se observa claramente como a medida que aumenta las horas de estudio aumenta el número de aprobados”.

Serían parcialmente correctas las estrategias de los estudiantes que funcionan para algunos problemas particulares, pero no son siempre válidas. Por ejemplo, se puede utilizar todos los datos del problema, pero realizar comparaciones aditivas, o bien trabajar con frecuencias absolutas y no con relativas, de modo que no se cuantifican adecuadamente las probabilidades como la dada en el problema 1 por el siguiente estudiante:

“Observando la tabla se deduce que 150 personas tienen insomnio, y de ellos 90 padecen estrés; además 100 personas no lo tienen y solo 60 de ellos padece estrés. Por lo que se puede deducir según las probabilidades que si puede tener relación”.

Cualquier otra estrategia sería incorrecta. Entre la variedad de estrategias incorrectas encontradas las más frecuentes son utilizar una sola celda, que es considerada por Inhelder y Piaget (1955) como propia del periodo de operaciones concretas y no de las operaciones formales. Generalmente se utiliza la celda que corresponde a presencia de dos caracteres en la tabla 2x2 o la celda de mayor frecuencia en la tabla 2x3, como el siguiente caso (ítem 4): *“100 personas: 51 aprobaron con más de 10 horas de estudio”.* Ello es consistente con los resultados de Inhelder y Piaget y de autores posteriores que analizaron estas estrategias (por ejemplo, Jenkins y Ward, 1965).

En la Tabla 3.8.2 se presentan los valores medios y error típico del valor estimado del coeficiente de asociación, en función del tipo de estrategia utilizada (correcta, parcialmente correcta o incorrecta). En el ítem 1 es donde aparece más clara la influencia de la estrategia correcta sobre la precisión, pues recordemos que en este ítem el coeficiente de asociación es cero, y el valor medio estimado disminuye notablemente con la corrección de la estrategia. En los otros ítems la influencia no aparece tan clara.

Para analizar si las diferencias observadas son estadísticamente significativas, se decidió realizar un contraste de hipótesis. Una vez comprobada previamente que la homogeneidad de varianzas en las tres muestras y en cada ítem permitía llevar a cabo el Análisis de varianza, así como el resto de supuestos del método (independencia, normalidad aproximada de la variable) se procedió a realizarlo, presentándose en la tabla el valor p correspondiente a dicha prueba, para cada ítem y en la última columna el valor empírico del coeficiente de asociación.

Tabla 3.8.2 Estimación de la asociación según el tipo de estrategia utilizada

Ítem	Sig. Anova	E. incorrectas		E. p. correctas		E. correctas		Coeficiente asociación
		Media	Error típico	Media	Error típico	Media	Error típico	
1	0,000	0,536	0,012	0,432	0,024	0,174	0,038	0
2	0,048	0,728	0,013	0,717	0,020	0,810	0,025	-0,62
3	0,209	0,693	0,015	0,727	0,021	0,756	0,029	0,67
4	0,001	0,785	0,015	0,846	0,009	0,797	0,008	0,37

Dicho valor p resultó estadísticamente significativo en todos los ítems, excepto para el ítem 3, correspondiente a la asociación directa e intensa y en el cuál los estudiantes no tenían ideas previas al respecto. Como hemos visto la precisión de la estimación fue muy buena en este caso, por lo que los estudiantes, incluso con una estrategia no correcta llegaron a una estimación bastante precisa del coeficiente. Para el resto de los ítems, dado que el resultado fue estadísticamente significativo, se estudiaron las diferencias entre los tres valores medios del coeficiente, correspondientes a las estrategias correctas, parcialmente correctas e incorrectas. Para saber entre qué grupos existen las diferencias se realizaron las pruebas post hoc de Tukey obteniéndose los siguientes resultados:

- Todos los pares de diferencias fueron estadísticamente significativas (al nivel 0,001) en el ítem 1, donde observamos como el valor del coeficiente se acerca más a 0 (verdadero valor) conforme mejora la estrategia (cuando pasa de incorrecta, a parcialmente correcta y correcta). En consecuencia, hay una fuerte influencia en este ítem de la corrección de la estrategia para la detección de la independencia; mientras que los estudiantes que utilizaron estrategias incorrectas o parcialmente correctas se guían más por sus teorías previas y sobrestiman el coeficiente de asociación, los que usan estrategias correctas se guían más por los datos y estiman un coeficiente más ajustado al valor verdadero.

- No hay un patrón claro de influencia de las estrategias sobre la precisión de la estimación en los otros ítems: Por un lado, en el ítem 2 únicamente se diferencian las estrategias correctas del resto y ello tan sólo a nivel $p=0,043$. Pero, en este caso, las estrategias correctas llevan a una mayor sobrestimación del coeficiente, aunque no demasiado grande. Por otro, en el ítem 4 solamente hay diferencias entre las estrategias incorrectas y parcialmente correctas (a nivel $p=0,01$) y también en este caso una mejor estrategia lleva a una mayor sobrestimación.

En definitiva, en la identificación de la independencia (ítem 1) la corrección de la estrategia mejora dicha identificación, no mostrándose una clara influencia en el resto de los ítems.

3.9. TEORÍAS PREVIAS SOBRE LA ASOCIACIÓN

Como se ha indicado, la investigación previa, por ejemplo, los trabajos de Jennings, Amabile y Ross (1982), Wrigth y Murphy (1984), Alloy y Tabacnick (1984) y Meiser y Hewstone (2006), sugiere la influencia de las teorías previas respecto al tipo de asociación sobre los juicios de asociación y la estimación del coeficiente. Resultados similares fueron obtenidos por Estepa y se han hecho patentes en las anteriores secciones.

Para comprobar si los estudiantes utilizan estas teorías previas explícitamente en sus justificaciones, se han analizado sus respuestas, Observando algunos ejemplos de estas teorías, como en las siguientes respuestas:

- Ítem 1: *“El no dormir bien, el no descansar puede influir en la carga de estrés de una persona”* (Estudiante 42).
- Ítem 2: *“Pienso que sí hay una relación entre tener o no hermanos y ser un niño sin o con problemas porque aquellos niños que no tengan hermanos se crían con valores morales mucho más distintos que los que sí los tienen. Valores como la generosidad, la empatía, etc.”* (Estudiante 44).
- Ítem 3: *“El llevar una vida sedentaria no tiene nada o poca relación con padecer alergia”* ó *“...utilizando la lógica entre las dos variables”* (Estudiante 12).
- Ítem 4: *“A mayor tiempo de estudio, más posibilidades de aprobar y menos de suspender”* (Estudiante 21).

Tabla 3.9.1. Frecuencia (y porcentajes) de estudiantes que expresan teorías previas

Ítem	Expresan una teoría	No la expresan	No contestan	Coefficiente de asociación	Teorías previas
1	40 (9,7)	372 (89,8)	2 (0,5)	0	En contra
2	33 (8)	376 (90,8)	5 (1,2)	-0,62	A favor
3	32 (7,7)	377 (91,1)	5 (1,2)	0,67	No hay
4	24 (5,8)	381 (92)	9 (2,2)	0,37	A favor

En la Tabla 3.9.1 mostramos la frecuencia de los estudiantes que hacen explícitas estas teorías. Pocos estudiantes expresan teorías previas, aunque en el ítem 1 deben utilizarlas implícitamente, pues muchos dan un juicio de asociación entre las variables. Se puede observar que se presenta el mayor número de estudiantes que expresan una teoría previa en la el ítem 1 (con un 9,7%), que curiosamente es en la única pregunta que las teorías previas van en contra de los datos.

Los resultados son compatibles con los de Estepa pues en su trabajo fue también en el primer ítem donde se expresaron más las teorías previas (23,5%) siendo mayor la proporción en su caso, así como en el ítem 4 (10,3%) y menor en los otros dos ítems que en nuestro caso.

3.10. ANÁLISIS DE ESTRATEGIAS

Otro punto estudiado fueron las estrategias, que diversos autores definen en forma diferente:

- Ríos (1976) define las estrategias como “*plan o conjunto de normas que establece un jugador y que indica cual es la acción que debe realizar en cada ocasión que corresponda actuar, en función de la información que tenga de las acciones y resultados anteriores al momento de su intervención*” (p. 23).
- En el informe Cockcroft (Cockcroft, 1985) se describen como “*procedimientos que guían la elección de la destreza que debe emplearse o de los conocimientos a que se debe recurrir en cada etapa de la resolución de un problema o del desarrollo de una investigación*”, (p. 87).
- Jennings, Amabile y Ross (1982), utilizan el término “*estrategias de evaluación de la asociación*” (p. 23) para indicar el modo en que los sujetos deciden cual es el grado de covariación en los datos presentados los sujetos.
- Pérez Echeverría (1990), al analizar los juicios de asociación, habla de estrategia para indicar las operaciones que hacen los sujetos con las celdas de la tabla de la siguiente forma: “*la estrategia [A] consiste en fijarse exclusivamente si el número*

de veces en que las dos variables concurren (p,q) es más alto que el resto de los casos” (p. 128).

- Ortega Martínez (1991) también usa la palabra en el mismo sentido “Otra estrategia que pueden utilizar los sujetos el evaluar la relación entre dos variables consiste en comparar las frecuencias de las celdas a y c ” (p. 34).
- Shaklee (1983) interpreta la palabra estrategia en la forma siguiente: “Las personas pueden combinar esta información de acuerdo con una variedad de estrategias para alcanzar una decisión. Muchas de estas estrategias pueden llevar a una respuesta correcta para un problema o un conjunto de problemas” (p. 434).

En lo que sigue analizamos las estrategias, dándole a este término el mismo sentido empleado en las investigaciones previas sobre juicios de asociación, por ejemplo, las de Jennings, Amabile y Ross (1982), Shaklee (1983), Pérez Echeverría (1990), Ortega Martínez (1991) y Estepa (1993). Sobre dichas estrategias hemos realizado varios tipos de análisis:

- En primer lugar se estudia el número de celdas empleadas, el tipo de frecuencia utilizada y los niveles de elaboración definidos por Pérez Echeverría (1990).
- En segundo lugar realizamos un análisis semiótico de las estrategias más frecuentes encontradas en nuestra muestra, dentro de cada uno de los niveles anteriores y clasificándolas como correctas, parcialmente correctas e incorrectas, al igual que en Estepa (1993). El análisis se basa en nuestro marco teórico e identifica diversos objetos y procesos matemáticos que intervienen en las prácticas matemáticas ligadas con la resolución de los problemas propuestos, así como los conflictos semióticos de los estudiantes. Unas tablas resúmenes muestran los objetos matemáticos utilizados en las diferentes estrategias y sirven como resumen de las configuraciones cognitivas identificadas. La finalidad es mostrar la complejidad de la asociación, incluso cuando su estudio se aborde únicamente desde el punto de vista intuitivo y /o de estadística descriptiva.
- Finalmente se describen y clasifican los conflictos semióticos identificados a lo largo del análisis semiótico de las estrategias.

En lo que sigue se presentan los resultados. Puesto que, en los análisis anteriores, el comportamiento de las diferentes muestras fue homogéneo, no se discrimina la muestra en el análisis de estrategias.

3.10.1. NIVELES DE ELABORACIÓN DE ESTRATEGIAS

En este apartado se analizan los niveles de complejidad definidos por Pérez Echeverría (1990), quien describe los siguientes, dando los porcentajes de uso para tablas 2x2. La autora se basa en el número de celdas utilizadas y en la forma que se realiza la comparación:

- Nivel 1. Utilizar sólo una celda, generalmente la celda a (uso 3,7% en la investigación de Pérez Echeverría).
- Nivel 2. Comparar la celda a con b o c , es decir usar dos datos (30,56% en la investigación citada).
- Nivel 3. Comparar la celda a con b y c , es decir usar tres datos (15,74% en dicho trabajo).
- Nivel 4. Usar todos los datos pero sólo mediante comparaciones aditivas (37,96%).
- Nivel 5. Concluir la asociación por el uso de una relación multiplicativa entre las cuatro celdas (12,04%).

En la Tabla 3.10.1.1 mostramos las frecuencias (y porcentajes) de los estudiantes que emplean un número determinado de celdas, en cada uno de los ítems del cuestionario.

Tabla 3.10.1.1. Frecuencia (y porcentaje) de estudiantes según las celdas utilizadas

	Tablas 2x2			Tabla 2x3
	Ítem 1	Ítem 2	Ítem 3	Ítem 4
0 celdas	13 (3,1)	15 (3,6)	15 (3,6)	8 (1,9)
1 celdas	73 (17,6)	20 (4,8)	33 (8)	8 (1,9)
2 celdas	108 (26,1)	154 (37,2)	153 (37)	44 (10,6)
3 celdas	27 (6,5)	16 (3,9)	9 (2,2)	75 (18,2)
4 celdas	179 (43,3)	201 (48,6)	178 (42,9)	49 (11,8)
5 celdas	-	-	-	0 (0)
6 celdas	-	-	-	208 (50,3)
No responde	14 (3,4)	8 (1,9)	26 (6,3)	22 (5,3)
Total	414	414	414	414

Como se puede observar en la tabla, en los tres primeros ítems (correspondientes a tablas de contingencia 2x2) el número de celdas usadas puede variar entre cero y cuatro celdas, y en el cuarto ítem (correspondiente a una tabla de contingencia 2x3) entre cero y seis. En base a ello, y teniendo en cuenta también el estudio de Estepa (1993), consideramos los siguientes niveles que extienden los de Pérez Echeverría (1990):

- *Nivel 0: Cuando no se emplea ninguna celda.* En este caso los estudiantes están basándose únicamente en sus teorías previas (Chapman y Chapman, 1969; Jennings, Amabile y Ross, 1982; Alloy y Tabacnick, 1984), sin considerar la tabla de contingencia, e introduciendo nuevos elementos sugeridos por su experiencia personal (Nisbett y Ross, 1980).
- *Nivel 1: Cuando se usa una celda,* los estudiantes usan sólo una parte de los datos, empleando estrategias como centrar su conclusión en el valor de mayor frecuencia, ó en el valor de menor frecuencia. En otros casos se utiliza la celda *a* de la tabla de contingencia 2x2 (Smedlund, 1963; Beyth y Marom, 1982; Shaklee y Mins, 1982, Yates y Curley, 1986). En todos estos casos los estudiantes manifiestan la concepción local de la asociación descrita por Estepa (1993).
- *Nivel 2: En caso de utilizar dos celdas en las tablas 2x2 o tres en la tabla 2x3,* los estudiantes también se apoyan en sólo una parte de los datos (concepción local), pero al menos emplean estrategias como en basarse en los valores de una distribución absoluta condicional, basarse en una distribución relativa condicional, ó fijarse en los valores de máxima y mínima frecuencia (Estepa, 1993; Estepa y Batanero, 1995).
- *Nivel 3: Cuando se trabaja con tres celdas en la tabla 2x2 y entre 3 y 5 en la 2x3,* los estudiantes emplean estrategias como comparar el primer valor con su valor de la derecha y su valor de debajo (en tablas 2x2), ó centrarse en el crecimiento de una distribución absoluta condicional (en tablas 2x3). Por lo general la celda que suelen los estudiantes dejar aparte suele ser la correspondiente a la ausencia de las dos variables ó a la celda de menor frecuencia (Estepa, 1993; Estepa y Batanero, 1995).
- *Cuando se usan cuatro celdas en la tabla 2x2 o seis en la 2x3,* los estudiantes están utilizando toda la información del problema, empleando estrategias como comparar filas o columnas. Dentro de estos estudiantes habrá que diferenciar estrategias aditivas (Nivel 4) y multiplicativas (Nivel 5), según Pérez Echeverría (1990). Debemos destacar que, aunque solamente las estrategias multiplicativas que utilizan todas las celdas son las que pueden ser matemáticamente correctas, en algunas ocasiones las matemáticamente no correctas producen juicios de asociación correctos (Shaklee, 1983), por lo que en estos casos hemos considerado la estrategia parcialmente correcta.

Otro punto a destacar es el reconocimiento, por parte del estudiante, de que para

emitir un juicio de asociación correcto, es necesario utilizar las 4 celdas en las tablas 2x2, o bien las 6 celdas en las tablas 2x3. Pues bien, en las tablas 2x2 el uso de las 4 celdas oscila entre el 42,9% y el 48,6%, lo que nos dice que más de la mitad de los estudiantes de la muestra no tienen conciencia de que es necesario utilizar todas las celdas para emitir un juicio de asociación correcto, sin embargo la mitad de la muestra (50,3%) si reconoce este hecho. Cuando se exige que se usen todas las celdas mediante una estrategia multiplicativa los porcentajes anteriores disminuyen considerablemente, ya que varían entre 11,9% y 18,3% (tabla 3.10.1.3).

En la Tabla 3.10.1.1 se observa que lo más frecuente fue usar todas las celdas, seguido por emplear dos en las tablas 2x2 y tres en la 2x3. Estos resultados coinciden con los obtenidos por Estepa (Tabla 3.10.1.2), salvo en el caso de dependencia directa en tabla 2x2, en el cual la máxima frecuencia en dicho trabajo fue el uso de dos celdas. Nuestros datos se aproximan más a las frecuencias indicadas por Pérez Echeverría (1990) que a las de Estepa (1993), seguramente será debido a que la muestra de Pérez Echeverría (1990) eran estudiantes de Psicología de los cursos 4º y 5º y la de Estepa (1993) eran estudiantes de COU.

Tabla 3.10.1.2. Frecuencia (y porcentaje) de estudiantes según las celdas utilizadas en la investigación de Estepa (1993)

	Tablas 2x2			Tabla 2x3
	D. Directa	D. Inversa	Independencia	D. Directa
1 celdas	29 (13,6)	26 (12,2)	22 (10,3)	6 (2,8)
2 celdas	62 (29,2)	40 (18,8)	50 (23,6)	31 (14,5)
3 celdas	1 (0,5)	0 (0)	0 (0)	69 (32,4)
4 celdas	60 (28,1)	125 (58,6)	115 (54)	2 (1)
5 celdas	-	-	-	0 (0)
6 celdas	-	-	-	78 (36,6)
Frecuencias marginales	-	-	-	7 (3,3)
Otras	56 (26,3)	18 (8,5)	23 (10,8)	14 (6,6)
No responde	5 (2,3)	4 (1,9)	3 (1,4)	6 (2,8)
Total	213	213	213	213

Teniendo en cuenta los niveles definidos por Pérez Echeverría (1990), y que hemos extendido anteriormente, obtenemos los resultados de la Tabla 3.10.1.3. Lo más frecuente fue usar una distribución condicional (nivel 2), excepto en el ítem 4 en que lo más frecuente fueron las estrategias aditivas en todos los datos (nivel 4). El segundo nivel más frecuente fue el 4 (en tablas 2x2) y el 2 (en la tabla 2x3). Menos de un 20% de estudiantes alcanza el nivel 5, aunque algo más que lo señalado por Pérez Echeverría.

Tabla 3.10.1.3. Frecuencia (y porcentaje) de estudiantes según el nivel y el ítem

	Tablas 2x2			Tabla 2x3
	Ítem 1	Ítem 2	Ítem 3	Ítem 4
Nivel 0: No usa celdas	13 (3,1)	15 (3,6)	15 (3,6)	8 (1,9)
Nivel 1: Una celda	73 (17,6)	20 (4,8)	33 (8)	8 (1,9)
Nivel 2: Una distribución condicional o 2celdas	108 (26,1)	154 (37,2)	153 (37)	119 (28,8)
Nivel 3: Más de una distribución condicional	27 (6,5)	16 (3,9)	9 (2,2)	49 (11,8)
Nivel 4: Todas las celdas, aditiva	103 (25)	127 (30,7)	119 (28,7)	159 (38,4)
Nivel 5: Todas las celdas: multiplicativa	76 (18,3)	74 (17,9)	59 (14,2)	49 (11,9)
No responde	14 (3,4)	8 (1,9)	26 (6,3)	22 (5,3)
Total	414	414	414	414

En la Tabla 3.10.1.4 presentamos los niveles obtenidos en los ítems equivalentes en el trabajo de Batanero, Estepa y Godino (1996). Observamos que en su caso para los ítems 1 y 4, la mayoría de los sujetos están en el nivel 2, y para los ítems 2 y 3, en esta investigación aparece el porcentaje más alto en un nivel 4. En el nivel 5 los porcentajes de Estepa en todos los ítems, excepto el cuarto son superiores a los nuestros.

Tabla 3.10.1.4. Frecuencia (y porcentaje) de estudiantes según el nivel utilizado en la investigación de Estepa (1993)

	Ítem 1	Ítem 2	Ítem 3	Ítem 4
Nivel 0: No usa celdas	56(26.3)	18(8.5)	23(10.8)	21(9.9)
Nivel 1: Una celda	29(13.6)	26(13.6)	22(10.3)	6(2.8)
Nivel 2: Una distribución condicional o 2 celdas	62(29.1)	40(18.8)	50(23.5)	100(46.9)
Nivel 3: Más de una distribución condicional	1(0.5)	0	0	2(0.9)
Nivel 4: Todas las celdas, aditiva	15(7.0)	74(34.7)	70(32.3)	66(31.0)
Nivel 5: Todas las celdas: multiplicativa	45(21.1)	51(23.9)	45(21.1)	12(5.6)
No responde	5(2.3)	4(1.9)	3(1.4)	6(2.8)

Tabla 3.10.1.5. Frecuencia (y porcentaje) de estudiantes según uso de frecuencia absoluta o relativa en cada ítem

	Tablas 2x2			Tabla 2x3
	Ítem 1	Ítem 2	Ítem 3	Ítem 4
Frecuencia absoluta	306 (74)	313 (75,7)	306 (74)	320 (77,3)
Frecuencia relativa	80 (19,3)	78 (18,8)	68 (16,4)	64 (15,5)
Ninguna	13 (3,1)	15 (3,6)	15 (3,6)	8 (1,9)
No responde	15 (3,6)	8 (1,9)	25 (6)	22 (5,3)
Total	414	414	414	414

Para complementar el estudio, en la Tabla 3.10.1.5 mostramos las frecuencias (y porcentajes) de los estudiantes que emplean en su estrategia una frecuencia absoluta ó una frecuencia relativa, en cada uno de los ítems del cuestionario. Como se puede observar en la tabla anterior, en todos los ítems se utiliza por lo general las frecuencias absolutas, por lo que los estudiantes no habrían entendido el problema de la asociación como de comparación de probabilidades, razonamiento que se supone adquirido en la

etapa de operaciones formales (Inhelder y Piaget, 1955). El uso de frecuencias absolutas indicaría en el mejor de los casos una estrategia parcialmente correcta.

Por su parte Konold, Pollatsek, Well y Gagnon (1997) sugieren que al comparar conjuntos de datos algunos estudiantes se centran en las frecuencias absolutas y no en las relativas, incluso cuando las muestras tienen tamaños muy diferentes, indicando falta de comprensión del concepto de distribución. Los casos, en los que no se utilizan ninguno de estos tipos de frecuencia, corresponden a los alumnos que no usan ninguna celda (por el uso de teorías previas, por ejemplo).

3.10.2. METODOLOGÍA DEL ANÁLISIS SEMIÓTICO DE ESTRATEGIAS

Un segundo estudio, más detallado, fue el análisis semiótico de estrategias, basado en nuestro marco teórico, en el cuál se da gran importancia a la idea función semiótica (Godino, 2002b; Font, Godino y D'Amore, 2007; Godino, Batanero y Font, 2007; Font, Planas y Godino, 2010).

Dicho estudio semiótico se realiza sobre las respuestas de todos los estudiantes de la muestra a los ítems. Puesto que dichas respuestas son datos textuales, podemos utilizar el análisis de contenido, que se basa en la idea de que las unidades del texto pueden clasificarse en un número reducido de categorías (Weber, 1985; Krippendorff, 1991). Tratamos, a partir de dicho análisis de realizar inferencias sobre las prácticas matemáticas y los conflictos semióticos de los estudiantes, mediante la identificación sistemática y objetiva de las características específicas de sus respuestas (Ghiglione y Matalón, 1991). Seguimos la misma técnica empleada por otros investigadores de nuestro grupo de investigación, por ejemplo, Olivo (2008) o Arteaga (2011).

Para cada ítem se realizó un proceso de análisis de datos cualitativo, inductivo y cíclico, de las respuestas de todos los estudiantes de la muestra siguiendo los siguientes pasos:

1. Se comenzó con el sistema de categorías de estrategias identificadas por Estepa (1993), que se describió en el Capítulo 2. Este sistema se fue depurando, comprobando, primeramente su presencia en nuestra muestra (pues algunas de las categorías de dicho autor no aparecieron). También se depuraron algunas de sus estrategias dividiéndolas en dos o más, en los casos en que nos pareció relevante.
2. Para llegar a nuestras categorías y partiendo de la lista anterior de Estepa (1993), en primer lugar, revisamos repetidamente todas las respuestas de cada uno de los ítems, comprobando el ajuste o desajuste de las respuestas con la lista inicial, añadiendo o

combinando elementos de esta lista, hasta conseguir una categorización que tiene en cuenta los conceptos, procedimientos, problemas, lenguaje, propiedades y argumentos usados explícitamente o usados implícitamente. Se asignó un código a cada categoría, se depuraron las categorías y se codificaron los datos.

3. Para cada una de las categorías obtenidas seleccionamos una respuesta representativa de la misma que muestre con claridad el argumento seguido por el estudiante.
4. Se procedió a efectuar un análisis semiótico de esta respuesta típica, siguiendo la pauta marcada en Godino, Wilhelmi y Font (2008), y también ejemplificada en Olivo (2008) y Arteaga (2011). Una vez elegida la respuesta, se transcribe literalmente y se separa en unidades de contenido. En el análisis semiótico se identifican los objetos matemáticos a los que va haciendo referencia el estudiante en su respuesta escrita y se analiza su naturaleza (definición, lenguaje, propiedad, etc.). También tratamos de identificar los procesos matemáticos involucrados en la resolución y especialmente los conflictos semióticos.
5. Los resultados se presentan en forma de tablas con tres columnas: la primera indica la unidad de análisis; la segunda incluye el fragmento de respuesta que constituye la unidad de análisis y la última muestra los objetos, procesos y conflictos semióticos identificados.

3.10.3. CONFIGURACIONES COGNITIVAS EN LAS ESTRATEGIAS

Cada una de las respuestas típicas en nuestra clasificación de estrategias lleva asociada una configuración de objetos y procesos matemáticos. Puesto que el significado dado a algunos de ellos no coincide con el institucional, se trata de configuraciones cognitivas (Font y Godino, 2006; Godino, Batanero y Font, 2007). A continuación se analizan, clasificando las estrategias dentro de cada uno de los niveles de Pérez Echeverría (1990), recordando que nos encontramos en la segunda situación de Vergnaud (1990), en la que los estudiantes no tienen las competencias para un uso rápido de las estrategias, pues no se trata de estrategias aprendidas, sino intuitivas. En el análisis semiótico se incluye el análisis de cada problema la primera vez que aparece, no repitiéndolo en las siguientes, para no ser reiterativos.

Estrategias de nivel 0

Son los casos en que no se usan las celdas de la tabla y donde los estudiantes muestran alguna concepción incorrecta sobre la asociación. Hemos diferenciado dos, según se expliciten o no las teorías previas.

SI.0.1. *El estudiante solo considera sus teorías y no tiene en cuenta los datos.*

En esta respuesta se muestra el fenómeno de correlación ilusoria descrito por Chapman y Chapman (1969) y Murphy y Medin (1985). El estudiante no utiliza los datos del problema ni ningún otro objeto matemático, y por lo tanto, matemáticamente la estrategia es incorrecta. En el siguiente ejemplo, analizado en la Tabla 3.10.3.1 el estudiante se guía por sus teorías previas, construidas a partir de su experiencia personal en lugar de utilizar los datos, y por tanto tiene un conflicto semiótico (procedimental). No es claro que sea capaz de leer la tabla, identificar los datos del problema, comprender la pregunta, en consecuencia, de utilizar una serie de conceptos y propiedades matemáticos ligados a la tabla de contingencia.

“Porque la variable de tener insomnio conlleva a padecer estrés, y no tener insomnio no conlleva a tener estrés” (Estudiante 5, ítem 1).

Tabla 3.10.3.1. Análisis semiótico de la estrategia SI.0.1.

U	Expresión	Contenido									
1	<p>Se pretende estudiar si el sufrir insomnio tiene relación con los trastornos de estrés. En una muestra de 250 personas observadas se obtuvieron los siguientes resultados:</p> <table border="1" style="margin-left: 20px;"> <thead> <tr> <th></th> <th>Padece Estrés</th> <th>No padece Estrés</th> </tr> </thead> <tbody> <tr> <td>Tener insomnio</td> <td>90</td> <td>60</td> </tr> <tr> <td>No tener insomnio</td> <td>60</td> <td>40</td> </tr> </tbody> </table>		Padece Estrés	No padece Estrés	Tener insomnio	90	60	No tener insomnio	60	40	<ul style="list-style-type: none"> • Se espera que el estudiante comience con un <i>proceso de significación</i> para interpretar el texto y sus elementos y que realice un <i>proceso de descomposición</i> del texto en elementos. • Se debe identificar las dos variables del enunciado (concepto) y establecer una correspondencia (procedimiento) entre los códigos (padecer o no estrés; tener o no insomnio, lenguaje) y los valores de la variable (concepto) en pacientes imaginarios (elemento fenomenológico) • No es claro que el estudiante realice los procesos anteriores e identifique el total de la muestra de 250 personas (concepto), o las frecuencias dobles, pues no las utiliza en la respuesta
	Padece Estrés	No padece Estrés									
Tener insomnio	90	60									
No tener insomnio	60	40									
2	<p>Utilizando estos datos razona si ser o no noctámbulo tiene relación con padecer estrés.</p>	<ul style="list-style-type: none"> • Ha de interpretar la pregunta (proceso de significación). • Ha de identificar lo que pide el problema: estimar si hay asociación entre las variables (procedimiento y conceptos). • Se espera que identifique si existe o no relación entre noctámbulo y padecer estrés (procedimiento y propiedad). • Se espera que realice una serie de <i>procesos de significación y de particularización</i> de varios conceptos al problema particular 									
3	<p><i>“Porque la variable de tener insomnio conlleva a padecer estrés y no tener insomnio conlleva a no padecer estrés”</i></p>	<ul style="list-style-type: none"> • El estudiante no hace referencia explícita a los datos de la tabla. • Aparece un <i>conflicto</i>, pues se guía por sus teorías previas, sin tener en cuenta los datos. • Proceso de argumentación: realiza una justificación de su respuesta (si hay relación). 									

SI.0.2. *Otros procedimientos incorrectos.* En casos esporádicos se usan procedimientos no relacionados con las frecuencias de la tabla, o bien no se especifica con claridad el procedimiento. Un ejemplo se presenta a continuación y se analiza en la Tabla 3.10.3.2. El estudiante considera en este ejemplo que: "si hay relación". Como en el caso anterior no es claro si ha sido capaz de leer la tabla, identificar los datos del problema, comprender la pregunta, ni tampoco si ha usado o no sus teorías previas.

"Por la observación de la tabla y utilizando la lógica entre los datos que se estudian" (Estudiante 12, ítem 2)

Tabla 3.10.3.2. Análisis semiótico de la estrategia SI.0.2.

U	Expresión	Contenido									
1	<p>Un psicólogo infantil estudia si existe relación entre ser hijo único (sin hermanos) y ser un niño problemático. Para ello se han observado 250 niños obteniendo los resultados siguientes</p> <table border="1" data-bbox="268 869 699 1093"> <thead> <tr> <th></th> <th>Niño con problemas</th> <th>Niño sin problemas</th> </tr> </thead> <tbody> <tr> <td>Tiene Hermanos</td> <td>40</td> <td>100</td> </tr> <tr> <td>Hijo Único</td> <td>100</td> <td>10</td> </tr> </tbody> </table>		Niño con problemas	Niño sin problemas	Tiene Hermanos	40	100	Hijo Único	100	10	<ul style="list-style-type: none"> No es claro que el estudiante identifique las dos variables del enunciado (concepto) y establezca una correspondencia (procedimiento) entre los códigos (tener hermanos o no; ser o no un niño con problemas) y los valores de la variable (concepto) en pacientes imaginarios (elemento fenomenológico) o que identifique los datos.
	Niño con problemas	Niño sin problemas									
Tiene Hermanos	40	100									
Hijo Único	100	10									
2	<p>Utilizando estos datos razona si hay alguna relación entre ser un niño con problemas y tener o no hermanos.</p>	<ul style="list-style-type: none"> Se espera que interprete la pregunta (<i>proceso de significación</i>). Ha de identificar lo que pide el problema: estimar si hay asociación entre las variables (conceptos). Ha de identificar una relación entre tener hermanos o no, con ser o no un niño con problemas. Debe realizar una serie de <i>procesos de significación y de particularización</i> de varios conceptos al problema particular. 									
3	<p>"Por la observación de la tabla y utilizando la lógica entre los datos que se estudian".</p>	<ul style="list-style-type: none"> El estudiante no hace referencia explícita a los datos de la tabla. Aparece un <i>conflicto</i>, pues no tiene en cuenta los datos. No es claro que los identifique. 									

Estrategias de nivel 1

Son los casos en que se usa una sola celda, por lo que supone una concepción local de la asociación (Estepa, 1993). Hemos encontrado algunas variantes que se analizan a continuación, todas ellas incorrectas.

SI.1.1. *Utiliza la celda a. Es decir, compara los casos en que se presentan los dos caracteres conjuntamente.* Esta estrategia fue descrita por Inhelder y Piaget (1955), y encontrada en Estepa (1993). También ha sido descrita en las investigaciones con sujetos adultos (Smedlund, 1963; Beyth y Marom, 1982; Shaklee y Mins, 1982; Yates y

Curley, 1986). Solamente se ha utilizado cuando la máxima frecuencia doble correspondía a la celda *a*. En este caso, el estudiante pensará que hay relación positiva si el valor de esta celda es mayor que el de las otras tres y que no la hay en caso contrario. Un ejemplo, donde el estudiante da un juicio de asociación positivo se presenta a continuación y se analiza en la tabla 3.10.3.3. Este estudiante ha sido capaz de leer la tabla e identificar al menos la frecuencia doble de la celda *a*; también de calcular la frecuencia relativa de dicha celda. Aparece un conflicto al considerar que puede deducir la asociación con sólo una parte de los datos (concepción local, según Estepa, 1993).

“Simplemente observando los resultados y relacionándolos con la muestra. Tener insomnio y padecer estrés supone un 0,4% de la muestra” (Estudiante 111, ítem 1).

Tabla 3.10.3.3. Análisis semiótico de la estrategia SI.1.1.

U	Expresión	Contenido
3	<i>“Simplemente observando los resultados y relacionándolos con la muestra. Tener insomnio y padecer estrés supone un 0,4% de la muestra”</i>	<ul style="list-style-type: none"> • Identifica el significado de cada valor de la tabla (procedimiento y proceso de significación). • Se fija en la celda <i>a</i> para llegar a su conclusión. • <i>Conflicto</i>: deduce incorrectamente la asociación de una parte de los datos, mostrando una concepción local de la misma. • Calcula la frecuencia relativa doble de la celda <i>a</i> (procedimiento y concepto). Hay un error de cálculo. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

SI.1.2. *Uso único de la celda de mayor frecuencia* (diferente de *a*), estrategia también encontrada en Estepa (1993). El estudiante no utiliza toda la información, tan sólo la celda más sobresaliente por su frecuencia, mostrando una concepción local de la asociación descrita por Estepa (1993), pues trata de decidir sobre la asociación sin tener en cuenta todos los datos del problema. Un ejemplo, donde el estudiante considera que *si hay relación*, donde hay un conflicto consistente en usar frecuencias absolutas se presenta en lo que sigue. Sin embargo, este estudiante fue capaz de interpretar la tabla y los datos, usando correctamente una serie de conceptos, procedimientos y procesos, como son las variables, valores, frecuencias absolutas dobles ó frecuencias absolutas marginales. Diferencia que el valor 51 es la intersección de dos sucesos (aprobar y estudiar más de 10 horas), además de encontrar una relación entre el valor 100 (frecuencia marginal) y el valor 51 (frecuencia absoluta doble).

“100 personas → 51 aprobaron con más de 10 horas de estudio” (Estudiante 48, ítem 4).

Tabla 3.10.3.4. Análisis semiótico de la estrategia SI.1.2.

U	Expresión	Contenido																				
1	<p>La siguiente tabla nos indica el número de estudiantes que aprobaron o suspendieron un examen teniendo en cuenta el tiempo que cada estudiante dedicó a prepararlo</p> <table border="1"> <thead> <tr> <th></th> <th>Menos de 5 h</th> <th>5 a 10 h</th> <th>Más de 10 h</th> </tr> </thead> <tbody> <tr> <td>Aprobados</td> <td>5</td> <td>15</td> <td>51</td> </tr> <tr> <td>Suspensos</td> <td>20</td> <td>7</td> <td>2</td> </tr> <tr> <td>TOTAL</td> <td>25</td> <td>22</td> <td>53</td> </tr> </tbody> </table> <table border="1"> <tbody> <tr> <td>TOTAL</td> </tr> <tr> <td>71</td> </tr> <tr> <td>29</td> </tr> <tr> <td>100</td> </tr> </tbody> </table>		Menos de 5 h	5 a 10 h	Más de 10 h	Aprobados	5	15	51	Suspensos	20	7	2	TOTAL	25	22	53	TOTAL	71	29	100	<ul style="list-style-type: none"> El estudiante comienza con un <i>proceso de significación</i> para interpretar el texto y sus elementos. Realiza un <i>proceso de descomposición</i> del texto en elementos. Identifica las dos variables del enunciado (concepto) y establece una correspondencia (procedimiento) entre los códigos (aprobar o suspender; número de horas) y los valores de la variable (concepto) en estudiantes imaginarios (elemento fenomenológico). Identifica las frecuencias dobles, en la tabla (conceptos y procedimiento) realizando un <i>proceso de particularización</i>. Identifica los totales por filas y los totales por columna (concepto). Identifica el valor 100 como el total (concepto).
	Menos de 5 h	5 a 10 h	Más de 10 h																			
Aprobados	5	15	51																			
Suspensos	20	7	2																			
TOTAL	25	22	53																			
TOTAL																						
71																						
29																						
100																						
2	Utilizando estos datos razona si aprobar o suspender el examen tiene relación con el tiempo dedicado a prepararlo.	<ul style="list-style-type: none"> Ha de interpretar la pregunta e identificar lo que pide el problema: estimar si hay asociación entre las variables (conceptos). Ha de identificar una relación entre aprobar o suspender y número de horas. Realiza una serie de <i>procesos de significación</i> y de <i>particularización</i> de varios conceptos al problema particular 																				
3	“100 personas \rightarrow 51 aprobaron con más de 10 horas de estudio”	<ul style="list-style-type: none"> Se fija en el valor total de la tabla, es decir el número de sujetos presentes en el experimento (concepto). Identifica el valor superior de las frecuencias, las celdas centrales, o moda 51 (más de 10 horas) (concepto procedimiento y propiedad). Representa con una flecha la relación entre el valor 100 y el valor 51 (concepto; <i>proceso de representación</i>). Compara estos dos resultados para ver si el número de aprobados con más de diez horas representa un buen indicador (procedimiento). <i>Conflicto</i>: deduce incorrectamente la asociación de una parte de los datos, mostrando una concepción local de la misma Otro <i>conflicto</i> es utilizar frecuencias absolutas en lugar de relativas. Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas. 																				

SI.1.3. *El uso de las celdas de menor frecuencia*. En este caso el estudiante sólo utiliza la celda menos sobresaliente, por lo que es una estrategia incorrecta y muestra la concepción local de la asociación (Estepa, 1993). Un ejemplo se presenta a continuación y se analiza en la Tabla 3.10.3.5. El estudiante considera que: “*si hay relación*” puesto que la celda de mínima frecuencia contradice la asociación, aplicando, por tanto, un razonamiento por contradicción. En este ejemplo, la estrategia, aunque incorrecta produce un juicio adecuado. El estudiante utiliza frecuencias absolutas dobles y marginales sin confundirlas, además de interpretarlas de forma correcta, reconociendo

una relación entre ellas al decir “2 de 53”. La hemos considerado sin embargo incorrecta por contener dos conflictos: (a) usar solo parte de los datos y (b) usar frecuencias absolutas.

“Si ya que según la tabla a más de 10 h de estudio menos probabilidad de suspender solo 2 de 53”. (Estudiante 60, ítem 4).

Tabla 3.10.3.5. Análisis semiótico de la estrategia SI.1.3.

U	Expresión	Contenido
3	“Si ya que según la tabla a más de 10 h de estudio menos probabilidad de suspender solo 2 de 53”.	<ul style="list-style-type: none"> • Identifica el valor menor de las frecuencias dobles (particularización de concepto, propiedad y procedimiento). • Compara con el total de la fila (procedimiento y concepto) • Asume que como este mínimo contradice la independencia, hay dependencia. • El <i>conflicto</i> aparece al no usar todos los datos del enunciado. • Otro <i>conflicto</i> es trabajar con frecuencias absolutas. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

SI.1.4. Se fija en alguna de las celdas que, a su juicio, contradicen la asociación directa, y si hay muchos casos en esta celda, supone que esto implica la no existencia de asociación. Esta estrategia es también descrita por Nisbett y Ross (1980) y Estepa (1993), quienes describen el uso de las celdas *b*, *c* y *d* como estrategias diferenciadas. Sin embargo, estos autores no hacen mención explícita a la búsqueda de la celda que, según el estudiante, contradice la asociación. Es una estrategia incorrecta, pues se basa en frecuencias absolutas. Un ejemplo, donde el estudiante indica “no asociación” se presenta a continuación y se analiza en la Tabla 3.10.3.6. El estudiante reconoce que los valores de la tabla (las frecuencias absolutas dobles) son la intersección de dos sucesos;

“No tiene relación ya que 60 personas que no tienen insomnio padecen estrés y es un porcentaje alto” (Estudiante 51, ítem 1).

Tabla 3.10.3.6. Análisis semiótico de la estrategia SI.1.4.

U	Expresión	Contenido
3	“No tiene relación ya que 60 personas que no tienen insomnio padecen estrés y es un porcentaje alto”	<ul style="list-style-type: none"> • Se fija en la celda <i>c</i> (presencia de una variable y ausencia de la otra). Esta celda contradice la asociación directa. (procedimiento, propiedad, concepto) • Identifica en la celda un valor elevado (no es el máximo). Particularización de una propiedad • <i>Conflicto</i> al asumir incorrectamente que si este valor es elevado implica independencia. El estudiante muestra una concepción determinista de la asociación. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

SI.1.5. Otros procedimientos. Algunos estudiantes usan procedimientos que no especifican con claridad, aunque se ve claro el uso único de una celda. Un ejemplo se

presenta a continuación y se analiza en la Tabla 3.10.3.7.

“Por medio de probabilidad se puede concluir que no influye, porque si influyera, los que no padecen insomnio deberían no padecer estrés, o padecer menos cantidad de sujetos estrés, cosa que no ocurre” (Estudiante 367, ítem 1).

Tabla 3.10.3.7. Análisis semiótico de la estrategia SI.1.5.

U	Expresión	Contenido
3	“Por medio de probabilidad se puede concluir que no influye, porque si influyera, los que no padecen insomnio deberían no padecer estrés, o padecer menos cantidad de sujetos estrés, cosa que no ocurre”	<ul style="list-style-type: none"> • Utiliza probabilidades, sin dejar claro qué tipo de probabilidad (concepto y procedimiento). • Como resultado concluye que no hay relación (procedimiento). • <i>Conflicto</i>: Considera que si hubiera relación entonces los que no padecen insomnio deberían no padecer estrés (concepción unidireccional) • <i>Conflicto</i> al usar únicamente la celda d (ausencia de los dos caracteres). • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

Estrategias de nivel 2

Se considera este nivel cuando los estudiantes usan 2 celdas para dar el juicio de asociación en la tabla 2x2 o tres celdas en la tabla 2x3. Generalmente se trata de estrategias incorrectas y son las siguientes:

SI.2.1. *Usar sólo una distribución condicional*. En este caso el estudiante no ve el problema como un problema de comparación de probabilidades, mostrando igualmente una concepción local de la asociación. En las investigaciones de Smedlund (1963) y Shaklee y Mins (1982), aparecen casos en los se comparaban la celda a (presencia de los dos caracteres) con la celda b , es decir utilizaban una distribución condicional. Está estrategia la vamos a dividir en dos tipos para los casos de tablas 2x2:

- Cuando utilizan en la distribución condicional la celda a , es decir la que indica la presencia de los dos caracteres (SI.2.1.1). Cuando no utilizan esta celda (SI.2.1.2). Separamos este caso por la importancia dada a esta celda en muchos estudios sobre asociación (como, por ejemplo, Smedlund, 1963; Beyth y Marom, 1982; Shaklee y Mins, 1982, Yates y Curley, 1986).

Un ejemplo del primer caso en una tabla 2x3 es el siguiente, donde el estudiante solo usa una fila y además estudia el crecimiento de las frecuencias a lo largo de los valores de esa fila (ver análisis en la Tabla 3.10.3.8). Es uno de los pocos casos en que la estrategia de este nivel es parcialmente correcta, El estudiante considera en este ejemplo que: “*si hay relación*”. La estrategia fallaría si los datos cambiasen, pues no se usan todos los datos. No hemos encontrado antecedentes de estrategias que impliquen usar el crecimiento o decrecimiento para estudiar la asociación. El

estudiante utiliza frecuencias absolutas dobles y reconoce las distribuciones condicionales absolutas, que son con las que trabaja.

“Simplemente he observado los datos de la tabla y según más horas se les dedique de estudio existe un mayor número de aprobados” (Estudiante 14, ítem 4).

Tabla 3.10.3.8. Análisis semiótico de la estrategia SI.2.1.

U	Expresión	Contenido
3	<i>“Simplemente he observado los datos de la tabla y según más horas se les dedique de estudio existe un mayor número de aprobados”.</i>	<ul style="list-style-type: none"> • Identifica los valores de las frecuencias absolutas de una de las distribuciones condicionales por filas, la del número de aprobados (concepto y procedimiento). • Compara los valores de las tres celdas que aparecen de esta variable en la tabla (procedimiento). • Identifica un crecimiento de la distribución según el tiempo (procedimiento; particularización de una propiedad). • Supone que ello implica la asociación. El <i>conflicto</i> se produce al usar sólo parte de los datos (concepción local) y otro <i>conflicto</i> por utilizar frecuencias absolutas • Aplica un argumento deductivo

SI.2.2. *Compara las frecuencias en una de las diagonales 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 estrategia también fue encontrada por Estepa (1993). En el ejemplo analizado en la Tabla 3.10.3.9 se utiliza la diagonal principal, comparando los valores de las dos celdas. El estudiante utiliza las frecuencias absolutas dobles y reconoce que las diagonales podrían ayudarle a realizar la interpretación (mediante la utilización de las celdas que informan de cada tipo de dependencia). Además reconoce el número total de sujetos del experimento en la tabla (250). Un conflicto se produce al no utilizar todos los datos y al considerar que la celda de no padecer ni estrés ni insomnio va en contra de la asociación.

“Padecer estrés – Tener insomnio > No padecer estrés – No insomnio; 90 > 40

Total 250 → 125

90 es el mayor número de sujetos con estrés e insomnio aunque no es una relación muy significativa” (Estudiante 61, ítem 1).

Tabla 3.10.3.9. Análisis semiótico de la estrategia SI.2.2.

U	Expresión	Contenido
3	<i>“Padecer estrés – Tener insomnio > No padecer estrés – No insomnio 90 > 40”</i>	<ul style="list-style-type: none"> • Utiliza símbolo de mayor > (concepto y representación). • Compara verbalmente los valores de las dos celdas que aparecen en la diagonal principal (procedimiento). • Compara numéricamente los valores de las dos celdas en la diagonal principal (procedimiento; propiedad del orden numérico).
4	<i>“Total 250 → 125”</i>	<ul style="list-style-type: none"> • Utiliza símbolo de flecha para establecer una relación (concepto y representación). • Divide mentalmente 250 entre dos (procedimiento y concepto), calculando la mitad del total de sujetos del experimento.

5	“90 es el mayor número de sujetos con estrés e insomnio aunque no es una relación muy significativa.”	<ul style="list-style-type: none"> • Asume incorrectamente que al ser mayor la frecuencia de sujetos con los caracteres que sin ellos hay dependencia (<i>conflicto</i> al aplicar una propiedad). • Un segundo <i>conflicto</i> aparece al no usar todos los datos del enunciado. • Otro <i>conflicto</i> es trabajar con frecuencias absolutas. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.
---	---	---

SI.2.3. *Uso de las celdas de mayor y menor frecuencia, mencionándolo explícitamente.* Tampoco estos estudiantes utilizan toda la información presente en el problema, sino tan sólo las celdas más sobresalientes. Estepa (1993) obtiene una estrategia consistente en fijarse sólo en la celda de mayor frecuencia, pero no el caso que describimos. Un ejemplo, donde se argumenta la asociación se presenta a continuación y se analiza en la Tabla 3.10.3.10. El estudiante emplea frecuencias absolutas dobles, identificando la mayor y menor; asume que la diferencia corresponde a la asociación.

“Primero he buscado el número mayor entre las relaciones y se corresponde a 51 (Más de 10 horas-aprobados) y el número más bajo 2 (más de 10 horas-suspensos)” (Estudiante 1, ítem 4).

Tabla 3.10.3.10. Análisis semiótico de la estrategia SI.2.3.

U	Expresión	Contenido
3	“Primero he buscado el número mayor entre las relaciones y se corresponde a 51 (Más de 10 horas-aprobados) y el número más bajo 2 (más de 10 horas-suspensos)”	<ul style="list-style-type: none"> • Identifica el valor mayor de las frecuencias dobles (particularización de propiedad y procedimiento; concepto). • Identifica el valor menor de las frecuencias dobles (particularización de propiedad y procedimiento; concepto). • Compara estas dos frecuencias absolutas (procedimiento y concepto). • Asume que la diferencia implica relación. El <i>conflicto</i> aparece al usar frecuencias absolutas • Otro <i>conflicto</i> aparece al no usar todos los datos del enunciado. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

SI.2.4. *Otros procedimientos usando dos o tres celdas en tablas 2x3.* En casos esporádicos se usan procedimientos que no se especifican suficientemente, aunque se observa el uso único de una distribución condicional. Un ejemplo se presenta a continuación (Tabla 3.10.3.11) donde se apoya la existencia de *relación*. Se producen los conflictos citados anteriormente. Solamente se ha utilizado en tablas 2x3.

“He concluido que si los estudiantes suspensos se pasan de la mitad de su total es que si tiene *relación*” (Estudiante 163, ítem 4).

Tabla 3.10.3.11. Análisis semiótico de la estrategia SI.2.4.

U	Expresión	Contenido
3	“He concluido que si los estudiantes suspensos se pasan de la mitad de su total es que si tiene relación”	<ul style="list-style-type: none"> • Identifica el total de estudiantes suspensos (particularización de propiedad y procedimiento). • Calcula la mitad de los estudiantes suspensos (procedimiento). • El <i>conflicto</i> aparece al no usar todos los datos del enunciado. • Otro <i>conflicto</i> es trabajar con frecuencias absolutas. • <i>Conflicto</i> al considera que si los estudiantes suspensos se pasan de la mitad de su total, entonces si tiene relación (procedimiento). • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

Estrategias de nivel 3

Se considera este nivel cuando usan tres celdas en la tabla 2x2 y más de tres celdas en la tabla 2x3, pues se usa más de una distribución condicional. Se han encontrado estrategias parcialmente correctas e incorrectas.

Estrategias parcialmente correctas

SP.3.1. *Comparar dos distribuciones condicionales por columnas, ignorando una distribución condicional* en tablas 2x3. Estepa (1993) encuentra un estudiante con esta estrategia. En este caso, el estudiante intenta transformar su tabla 2x3 a una tabla 2x2, y por tanto, es una variante de la estrategia “usar dos distribuciones condicionales en tablas 2x2”. Es parcialmente correcta, porque pudiera ser que las dos distribuciones comparadas fuesen idénticas y la que falta por comparar diferente de las anteriores. Sin embargo, en algunos casos da resultados correctos si se comparan dos distribuciones diferentes. Un ejemplo se presenta a continuación (Tabla 3.10.3.12) donde el estudiante utiliza frecuencias absolutas dobles y distribuciones condicionales absolutas. Con un procedimiento informal, correcto, se centra en las celdas que informan de dependencia directa e inversa, aunque este procedimiento es sólo válido para tablas 2x2.

“He llegado a esa conclusión porque cuando se le dedica menos tiempo al estudio hay mayor número de suspensos y cuando le dedicas mayor tiempo hay el triple de aprobados.” (Estudiante 67, ítem 4).

Tabla 3.10.3.12. Análisis semiótico de la estrategia SI.3.1.

U	Expresión	Contenido
3	“He llegado a esa conclusión porque cuando se le dedica menos tiempo al estudio hay mayor número de suspensos y cuando le dedicas mayor tiempo hay el triple de aprobados.”	<ul style="list-style-type: none"> • Identifica la primera y última distribución condicional por columnas (concepto). • El estudiante compara los valores de la distribución condicional en estas dos columna, buscando el valor mayor (procedimiento). • El estudiante establece una proporción 1:3 (concepto; procedimiento). • El <i>conflicto</i> aparece al trabajar con frecuencias absolutas. Otro <i>conflicto</i> es que no usa todas las columnas. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

Estrategias incorrectas

SI.3.2. *Compara la celda a (presencia de los dos caracteres) con las celdas b y c (presencia de uno de los dos caracteres).* Sólo aplicable a tablas 2x2. Es decir, comparan los casos en que presentan los dos caracteres conjuntamente, con aquellos en que se presenta sólo uno de los dos. Es una variante de la estrategia “usar la celda a” descrita por Estepa (1993) y Pérez Echeverría (1990) y es incorrecta. Un ejemplo, donde el estudiante considera que “*hay relación*” se presenta a continuación y se analiza en la Tabla 3.10.3.13. El estudiante usa frecuencias absolutas dobles y frecuencias marginales y, reconoce las celdas que informan o contradicen la asociación, pero presenta varios conflictos.

“Pues he observado que hay más personas en la muestra que padecen estrés y además tienen insomnio (90 personas) y que exactamente el mismo número de personas dentro de los subconjuntos de no tener insomnio y padecer estrés y no padecer estrés y tener insomnio” (Estudiante 151, ítem 4).

Tabla 3.10.3.13. Análisis semiótico de la estrategia SI.3.2.

U	Expresión	Contenido
3	<i>“Pues he observado que hay más personas en la muestra que padecen estrés y además tienen insomnio (90 personas) y que exactamente el mismo número de personas dentro de los subconjuntos de no tener insomnio y padecer estrés y no padecer estrés y tener insomnio”</i>	<ul style="list-style-type: none"> • Reconoce el valor mayor de la tabla (procedimiento). • Reconoce la igualdad de los valores en una de las diagonales (procedimiento). • Asume incorrectamente que al cumplirse los puntos anteriores, implica que existe una relación entre las variables (<i>conflicto</i> al aplicar una propiedad). • El <i>conflicto</i> aparece al no usar todos los datos del enunciado. • Otro <i>conflicto</i> es trabajar con frecuencias absolutas. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

SI.3.3 (Sólo aplicable a tablas 2x2). *Compara la celda a (presencia de los dos caracteres) con las celdas c (sólo un carácter) y d (ausencia de los dos caracteres) ó compara la celda a con las celdas b y d.* Es decir, comparan los casos en que presentan los dos caracteres conjuntamente, con aquellos en que se presenta sólo uno de los dos y en los que no se presenta ninguno. Es una variante de la estrategia “usar la celda a” descrita por Estepa (1993), quien no la cita explícitamente y es incorrecta. En el ejemplo analizado en la Tabla 3.10.3.14, el estudiante muestra dos conflictos semiótico: uno de ellos basar sus comparaciones en frecuencias absolutas y otro usar sólo una parte de los datos, es decir, mostrar la concepción local de la asociación (Estepa, 1993).

“Al igual que en el estudio anterior, creo que si porque hay una mayoría de niños con problemas siendo hijo único, que teniendo hermanos, existiendo una diferencia de 60 sujetos.”
“Y por otro lado, los niños con problemas respecto a los que no lo tienen se distancian entre 90 participantes, siendo los que si padecen éstos 100 sujetos.” (Estudiante 52, ítem 3).

Tabla 3.10.3.14. Análisis semiótico de la estrategia SI.3.3.

U	Expresión	Contenido
3	“Al igual que en el estudio anterior, creo que si porque hay una mayoría de niños con problemas siendo hijo único, que teniendo hermanos, existiendo una diferencia de 60 sujetos.”	<ul style="list-style-type: none"> • El estudiante establece una relación en la estrategia con el ejercicio anterior (procedimiento). • El estudiante se centra en los niños hijos únicos con problemas (particularización). • Compara los niños únicos-problemáticos, con los niños con hermanos-problemáticos, es decir compara los valores de una distribución condicional por columna (procedimiento). • Realiza operación de restar estos valores (procedimiento).
4	“Y por otro lado, los niños con problemas respecto a los que no lo tienen se distancian entre 90 participantes, siendo los que si padecen éstos 100 sujetos.”	<ul style="list-style-type: none"> • Compara los niños únicos-problemáticos, con los niños únicos-no problemáticos, es decir compara los valores de una distribución condicional por fila (procedimiento). • Realiza operación de restar estos valores (procedimiento). • El <i>conflicto</i> aparece al no usar todos los datos del enunciado. • Otro <i>conflicto</i> es trabajar con frecuencias absolutas. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

SI.3.4. *Compara la celda d (ausencia de los dos caracteres) con las celdas b y c* (sólo aplicable a tablas 2x2). Es decir, comparan los casos en que no se presenta ninguno de los dos caracteres conjuntamente, con aquellos en que se presenta sólo uno de los dos. No hemos encontrado esta estrategia en los antecedentes, aunque Estepa encuentra estudiantes que usan la celda *d* o las celdas *b* y *c*. Un ejemplo se presenta a continuación (Tabla 3.10.3.15). Aparecen los conflictos señalados anteriormente.

“Si en una muestra de 250 niños, 100 niños son hijos únicos problemáticos y sólo 10 que no son problemáticos y se compara con que 100 niños que tienen hermanos no tienen problemas me da a entender que ser hijo único de la probabilidad de ser mas problemático que uno que tiene hermanos.” (Estudiante 252, ítem 2).

Tabla 3.10.3.15. Análisis semiótico de la estrategia SI.3.4.

U	Expresión	Contenido
3	“Si en una muestra de 250 niños, 100 niños son hijos únicos problemáticos y sólo 10 que no son problemáticos y se compara con que 100 niños que tienen hermanos no tienen problemas me da a entender que ser hijo único de la probabilidad de ser mas problemático que uno que tiene hermanos.”	<ul style="list-style-type: none"> • El estudiante identifica el total de la muestra de 250 niños mencionado (concepto y procedimiento). • Compara los niños únicos-problemáticos, con los únicos-no problemáticos, es decir compara los valores de una distribución condicional por fila (concepto y procedimiento). • Compara los niños únicos-problemáticos, con los niños con hermanos-no problemáticos, es decir compara los valores de una diagonal (procedimiento). • El <i>conflicto</i> aparece al no usar todos los datos del enunciado. • Otro <i>conflicto</i> es trabajar con frecuencias absolutas. Confunde entonces probabilidades con casos posibles. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

SI.3.5. *Otros procedimientos basados en más de una distribución condicional.* En el caso de tabla 2x3 son procedimientos donde se utilizan 4 ó 5 celdas; en algunos casos se intenta emplear la probabilidad condicionada. En el siguiente ejemplo (Tabla 3.10.3.16) el estudiante realiza un proceso de representación e intenta calcular la

probabilidad condicional de haber aprobado si se ha estudiado más de 10 horas; pero desarrolla mal la fórmula, pues usa la frecuencia marginal (número de aprobados) en vez de la frecuencia doble (aprobados y que hayan estudiado más de 10 horas). Por un lado hay una confusión entre probabilidad simple y conjunta. Por otro, el estudiante no es consciente de que una probabilidad no puede ser mayor que la unidad; ambos errores fueron descritos por Contreras (2011).

“ $A = \text{aprobar} / A^C = \text{suspender}$
 $B = < 5h / C = 5-10h / D = > 10h$
 $P(A/D) = 71/53$ ” (Estudiante 160, ítem 4).

Tabla 3.10.3.16. Análisis semiótico de la estrategia SI.3.5.

U	Expresión	Contenido
3	$A = \text{aprobar} / A^C = \text{suspender}$ $B = < 5h / C = 5-10h /$ $D = > 10h$ $P(A/D) = 71/53$	<ul style="list-style-type: none"> • Asignación de letras a cada columna y a cada fila (procedimiento y proceso de representación). • Utiliza símbolo de igualdad, mayor y menor para establecer una relación (concepto y representación). • Trata de calcular la probabilidad condicionada (procedimiento y concepto). Pero tiene un <i>conflicto</i> al utilizar la frecuencia marginal (número de aprobados) en vez de la frecuencia doble (aprobados que han estudiado más de 10 horas). • <i>Conflicto</i> al confundir probabilidad simple y conjunta • <i>Conflicto</i> al no ser consciente de que una probabilidad no puede ser mayor que la unidad. • <i>Otro conflicto</i> es no usar todos los datos del enunciado. • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.

Estrategias de nivel 4

Se considera este nivel cuando el estudiante emplea cuatro celdas en la tabla 2x2 y seis celdas en la tabla 2x3, pero las combina mediante estrategias aditivas, por lo que en muchos casos tratarán de deducir la asociación a partir de frecuencias absolutas. Recordemos que autores como Godino y Batanero (1998) indican que la comparación de dos o más muestras tiene que realizarse en términos de las frecuencias relativas. Dentro de este grupo hemos encontrado estrategias parcialmente correctas e incorrectas.

Estrategias parcialmente correctas

SP.4.1. *Comparar la suma de los valores en las diagonales.* En este caso, descrito por Shaklee y Tucker (1980), Allan y Jenkins (1983) y Estepa (1993), el estudiante calcula la suma de las diagonales. En el ejemplo que se presenta en lo que sigue y se analiza en la Tabla 3.10.3.17, el estudiante utiliza todas las celdas. Emplea un procedimiento informal de forma correcta, reconociendo la relación entre las frecuencias absolutas de la diagonal principal con la asociación directa y comparando

estos datos con el número total de sujetos en el experimento (250). Por otro lado, repite el procedimiento con las celdas que corresponden a dependencia inversa. Pero no llega a utilizar estrategias multiplicativas.

*“Los niños que tienen hermanos y además son problemáticos son solo 40 de 250 niños y los niños que son hijos únicos sin problemas son solamente 10 de 250. Esto indica que están relacionados y que en general los hijos únicos son más problemáticos. Por otro lado también podemos decir que la mayoría de los hijos únicos, son niños con problemas (100 de 250) y la mayoría de los hijos con hermanos (100 de 250) son niños sin problemas. Hijos únicos => + problemas
Hijos con hermanos => - problemas”* (Estudiante 172, ítem 2).

Tabla 3.10.3.17. Análisis semiótico de la estrategia SP.4.1.

U	Expresión	Contenido
3	<p><i>“Los niños que tienen hermanos y además son problemáticos son solo 40 de 250 niños y los niños que son hijos únicos sin problemas son solamente 10 de 250. Esto indica que están relacionados y que en general los hijos únicos son más problemáticos. Por otro lado también podemos decir que la mayoría de los hijos únicos, son niños con problemas (100 de 250) y la mayoría de los hijos con hermanos (100 de 250) son niños sin problemas. Hijos únicos => + problemas Hijos con hermanos => - problemas”</i></p>	<ul style="list-style-type: none"> • El estudiante identifica el total de la muestra de 250 niños mencionado (concepto; proceso de particularización). • El estudiante se centra inicialmente en los valores de la diagonal principal. Establece una relación entre las frecuencias absolutas y el total de la muestra: 40 de 250, 10 de 250 y 100 de 250. Es el paso previo al establecimiento del tanto por ciento. Deduce que la segunda relación es mayor que la primera. (conceptos, procedimientos, particularización). • Repite el procedimiento con los valores de la otra diagonal (conceptos, procedimientos, particularización). • Compara las dos diagonales (procedimiento). El conflicto aparece al trabajar con frecuencias absolutas. • Utiliza símbolos de flechas y de + para establecer una relación (representación; procedimiento). • Aplica un argumento deductivo que contiene un proceso de síntesis.

SP.4.2. *Comparar dos distribuciones condicionales en forma aditiva.* En esta estrategia se utilizan todas las frecuencias de la tabla; pero no se relacionan entre sí en forma multiplicativa, sino aditiva. Fue encontrada en el estudio de Inhelder y Piaget (1955) y también aparece en la comparación de probabilidades. Autores como Estepa (1993) o Godino y Batanero (1998) nos indican la necesidad de usar frecuencias relativas al comparar dos distribuciones. Un ejemplo se presenta a continuación (ver análisis en la Tabla 3.10.3.18), donde el estudiante compara en cada una de las dos distribuciones condicionales las frecuencias absolutas de casos a favor y en contra de la variable. Como la diferencia es mayor en una de las distribuciones, asume que esto es sinónimo de asociación. La estrategia es parcialmente correcta porque podría funcionar si las frecuencias marginales hubiesen sido iguales. No obstante, manifiesta el conflicto consistente en comparar probabilidades en forma aditiva, además de los ya mencionados con anterioridad.

“Para llegar a la conclusión he mirado que hay 100 niños hijos únicos que tienen problemas y sólo 10 hijos únicos sin problemas, y después he mirado que hay 100 niños que tienen hermanos sin problemas y 40 niños con hermanos y con problemas. Por lo que existe una relación entre ser hijo único y tener problemas.” (Estudiante 67, ítem 2).

Tabla 3.10.3.18. Análisis semiótico de la estrategia SP.4.2.

U	Expresión	Contenido
3	“Para llegar a la conclusión he mirado que hay 100 niños hijos únicos que tienen problemas y sólo 10 hijos únicos sin problemas, y después he mirado que hay 100 niños que tienen hermanos sin problemas y 40 niños con hermanos y con problemas. Por lo que existe una relación entre ser hijo único y tener problemas.”	<ul style="list-style-type: none"> • El estudiante identifica la distribución condicionada a ser hijo único (concepto y procedimiento). • El estudiante identifica en esta distribución condicionada la frecuencia de niños con y sin problemas (concepto). • El estudiante compara estas dos frecuencias (concepto y procedimiento). • Repite los anteriores pasos en la distribución condicionada a tener hermanos (concepto y procedimiento). • Al ser mayor la diferencia en una de las distribuciones, supone asociación. Hay un <i>conflicto</i> al realizar comparaciones aditivas en lugar de multiplicativas. Otro <i>conflicto</i> aparece al trabajar con frecuencias absolutas • Aplica un argumento deductivo ligando todos los anteriores procesos y justificaciones.

SP.4.3. *Comparar las frecuencias marginales de cada distribución condicional, confundiendo los conceptos media y frecuencia.* En el ejemplo analizado en la Tabla 3.10.3.28, se comparan las frecuencias absolutas dobles en cada distribución condicional, considerando que al ser mayor la diferencia en una de las distribuciones hay asociación. Como ya hemos comentado en la estrategia anterior, es parcialmente correcta. Se añade en este ejemplo la confusión entre frecuencia y valor de la variable, un conflicto que fue citado por Cobo (2003) y Mayén (2009) al estudiar la comprensión de las medidas de tendencia central.

“He llegado a esta conclusión teniendo en cuenta que hay más hijos únicos con problemas (100) que hijos que tienen hermanos (40). Pero al calcular la media como hice en el ejercicio anterior he comprobado que hay más niños con hermanos y con problemas (70) que niños sin hermanos y sin problemas (55)” (Estudiante 40, ítem 2).

Tabla 3.10.3.28. Análisis semiótico de la estrategia SP.4.3.

U	Expresión	Contenido
3	“He llegado a esta conclusión teniendo en cuenta que hay más hijos únicos con problemas (100) que hijos que tienen hermanos (40). Pero al calcular la media como hice en el ejercicio anterior he comprobado que hay más niños con hermanos y con problemas (70) que niños sin hermanos y sin problemas (55)”.	<ul style="list-style-type: none"> • Ha comparado en la distribución condicional la frecuencia de tener hermanos, en niños con problemas y sin problemas (conceptos y procedimiento; particularización). • <i>Conflicto</i> al confundir los conceptos de frecuencia doble y media. • <i>Conflicto</i> al trabajar con frecuencias absolutas. • Compara las frecuencias dobles (procedimiento). Supone que el ser diferentes implica dependencia • Aplica un argumento deductivo

Estrategias incorrectas

SI.4.4. *Comparar la diferencia de los valores de las frecuencias en las diagonales en las tablas 2x2.* Estepa (1993) describe la comparación de la suma de diagonales, que es una estrategia parecida a esta, y en otras investigaciones se cita la comparación de las diagonales; por ejemplo, en Allan y Jenkins (1983) y Shaklee y Tucker (1980). Un ejemplo, donde el estudiante considera que hay relación, se presenta a continuación y se analiza en la Tabla 3.10.3.19. El estudiante reconoce la importancia de las diagonales en la detección de la asociación, ya que se observan las celdas que informan de dependencia directa y de la dependencia inversa, aunque el procedimiento que utiliza es erróneo, pues no es la diferencia de estos valores sino su suma la que informa de la asociación.

“Observando la poca variación entre 130 – 120 y 20 – 30 (sólo de 10)” (Estudiante 50, ítem 3).

Tabla 3.10.3.19. Análisis semiótico de la estrategia SI.4.4.

U	Expresión	Contenido									
1	<p>Se quiere saber si sufrir o no de alergia tiene relación con llevar una vida sedentaria (llevar una vida sin realizar ningún tipo de ejercicio físico). Para ello obtuvimos los siguientes resultados de 300 sujetos:</p> <table border="1" style="margin-left: 20px;"> <thead> <tr> <th></th> <th>Sufre alergia</th> <th>No sufre alergia</th> </tr> </thead> <tbody> <tr> <td>Vida Sedentaria</td> <td>130</td> <td>30</td> </tr> <tr> <td>Vida no sedentaria</td> <td>20</td> <td>120</td> </tr> </tbody> </table>		Sufre alergia	No sufre alergia	Vida Sedentaria	130	30	Vida no sedentaria	20	120	<ul style="list-style-type: none"> • El estudiante comienza con un <i>proceso de significación</i> para interpretar el texto y sus elementos. Realiza un <i>proceso de descomposición</i> del texto en elementos. • Identifica las dos variables del enunciado (concepto) y establece una correspondencia (procedimiento) entre los códigos (forma de vida sedentaria o no; sufrir alergia o no) y los valores de la variable (concepto) en pacientes imaginarios (elemento fenomenológico), además del total de la muestra de 300 sujetos (concepto). • Identifica las frecuencias dobles, en la tabla (conceptos y procedimiento) realizando un <i>proceso de particularización</i>.
	Sufre alergia	No sufre alergia									
Vida Sedentaria	130	30									
Vida no sedentaria	20	120									
2	<p>Utilizando estos datos razona si padecer alergia tiene o no relación con llevar una vida sedentaria en estos sujetos.</p>	<ul style="list-style-type: none"> • Ha de interpretar la pregunta e identificar lo que pide el problema: estimar si hay asociación entre las variables (conceptos). • Ha de identificar una relación entre tener forma de vida sedentaria o no, con sufrir alergia o no. • Realiza una serie de procesos de significación y de particularización de varios conceptos al problema particular 									
3	<p>“Observando la poca variación entre 130 – 120 y 20 – 30 (sólo de 10)”</p>	<ul style="list-style-type: none"> • Identifica los casos favorables y desfavorables a la asociación y trata de hallar una relación en las diagonales (particularización de una propiedad). Calcula la diferencia de frecuencias en la diagonal que apoya la relación (procedimiento). • <i>Conflicto</i>: Calcula la diferencia de frecuencias la diagonal que no apoya la relación (procedimiento). Asume incorrectamente que esta pequeña diferencia es indicativa de no asociación. • Otro <i>conflicto</i> es trabajar con frecuencias absolutas. • Aplica un argumento deductivo 									

SI.4.5. *Comparar todas las frecuencias absolutas dobles entre si.* Esta estrategia también fue encontrada en Estepa (1993) y, como en casos anteriores, supone falta de comprensión de la necesidad de trabajar con frecuencias relativas al comparar distribuciones. En el siguiente ejemplo, analizado en la Tabla 3.10.3.20, la argumentación es muy confusa, pues el estudiante se limita a comparar frecuencias absolutas, sin llegar a un juicio claro de asociación.

“Porque tener insomnio/padecer estrés tiene una puntuación muy alta (90). De todos modos no es una relación muy alta porque no tener insomnio/padecer estrés y tener insomnio /no padecer estrés también es bastante alta (60) y no tener insomnio/no padecer estrés es bastante más baja (40).La conclusión es que padecer estrés si que es una razón con mucho peso para tener insomnio (o viceversa) pero hay alguna razón con bastante peso también” (Estudiante 156, ítem 1).

Tabla 3.10.3.20. Análisis semiótico de la estrategia SI.4.5.

U	Expresión	Contenido
3	<i>“Porque tener insomnio/padecer estrés tiene una puntuación muy alta (90). De todos modos <u>no</u> es una relación muy alta porque no tener insomnio/padecer estrés y tener insomnio /no padecer estrés también es bastante alta (60) y no tener insomnio/no padecer estrés es bastante más baja (40).”</i>	<ul style="list-style-type: none"> • Identifica la celda donde se dan los dos síntomas: insomnio y estrés y observa que es la de máxima frecuencia (concepto y procedimiento). • Identifica las celdas en las que se da uno de los síntomas y compara con la anterior (procedimiento). • El estudiante considera que no es una relación muy alta porque en las celdas en las que se da uno de los dos caracteres es bastante alto (propiedad). • <i>Conflicto</i>: llamar a las frecuencias puntuación. • <i>Conflicto</i>: trabajar con frecuencias absolutas. • Identifica las celdas en las que no se da ninguno de los síntomas y ve que la frecuencia es menor (concepto y procedimiento).
4	<i>“La conclusión es que padecer estrés si que es una razón con mucho peso para tener insomnio (o viceversa) pero hay alguna razón con bastante peso también.”</i>	<ul style="list-style-type: none"> • El estudiante compara las frecuencias absolutas dobles (procedimiento). • El <i>conflicto</i> aparece al trabajar con frecuencias absolutas. • <i>Conflicto</i>, al realizar comparaciones aditivas • Aplica un argumento deductivo , pero su argumentación es muy confusa

SI.4.6. *Comparar los totales de las distribuciones condicionadas.* Es una estrategia incorrecta, pues sólo usa una de las dos variables. En el estudio de Estepa (1993), esta estrategia sólo se encontró en tablas 2x3. En el ejemplo que se presenta a continuación (Tabla 3.10.3.21) el estudiante llega a esta conclusión de que *“si hay relación”* trabajando sólo con frecuencias marginales, lo que implica un conflicto, pues no se puede deducir la asociación del estudio aislado de una de las dos variables. Por otro lado, trabaja, como en estrategias anteriores con frecuencias absolutas.

“Si el total de los sujetos son 100, 71 de los mismos han aprobado y en cambio 29 han suspendido” (Estudiante 45, ítem 4).

Tabla 3.10.3.21. Análisis semiótico de la estrategia SI.4.6.

U	Expresión	Contenido
3	“Si el total de los sujetos son 100, 71 de los mismos han aprobado y en cambio 29 han suspendido”.	<ul style="list-style-type: none"> • Identifica la distribución marginal (concepto). • Sólo utiliza las frecuencias marginales, es decir los totales de aprobados y suspensos, y totales en el experimento (conceptos; particularización). • <i>Conflicto</i> al comparar frecuencias marginales (procedimiento). Supone que la asociación se deduce de una de las frecuencias marginales sin considerar la otra variable, en este caso “horas de estudio”. • Otro <i>conflicto</i> aparece al trabajar con frecuencias absolutas. • Aplica un argumento deductivo

SI.4.7. *Otros procedimientos de nivel 4.* En casos aislados los estudiantes usan otros procedimientos, empleando todos los datos. Un ejemplo se presenta a continuación (Tabla 3.10.3.22) donde el estudiante usa estrategias aditivas para comparar los cuatro valores de la tabla. Estas estrategias, según Inhelder y Piaget (1955), son todavía de transición entre el nivel de operaciones concretas y formales.

“La tabla nos indica que las personas que tienen insomnio y padecen estrés varían en una diferencia de 30 personas respecto a las que tienen insomnio pero no padecen estrés y otra diferencia de 20 personas en relación a las que no tienen insomnio pero pueden no padecer estrés. Por lo tanto, no hay una relación clara que nos indique que estas dos variables estén conectadas entre sí” (Estudiante 307, ítem 1).

Tabla 3.10.3.22. Análisis semiótico de la estrategia SI.4.7.

U	Expresión	Contenido
3	“La tabla nos indica que las personas que tienen insomnio y padecen estrés varían en una diferencia de 30 personas respecto a las que tienen insomnio pero no padecen estrés y otra diferencia de 20 personas en relación a las que no tienen insomnio pero pueden no padecer estrés. Por lo tanto, no hay una relación clara que nos indique que estas dos variables estén conectadas entre sí”.	<ul style="list-style-type: none"> • El estudiante identifica en la distribución condicionada insomnio la frecuencia de estudiantes con y sin estrés (concepto). • Calcula la diferencia de estas dos frecuencias (procedimiento). • Repite los anteriores pasos en la distribución condicionada a no tener insomnio (concepto y procedimiento). • Al ser las diferencias similares, supone que no hay relación. Hay un <i>conflicto</i> al realizar comparaciones aditivas en lugar de multiplicativas. Otro <i>conflicto</i> aparece al trabajar con frecuencias absolutas. • Aplica un argumento deductivo ligando todos los anteriores procesos y justificaciones.

Estrategias de nivel 5

Se considera este nivel cuando el estudiante usa todas las celdas de la tabla para establecer el juicio de asociación y las compara mediante operaciones multiplicativas. Hemos encontrado estrategias correctas, parcialmente correctas e incorrectas, que se describen en lo que sigue.

Estrategias correctas

Los estudiantes que las usan, se apoyan intuitivamente en propiedades de la asociación, correctamente identificadas y aplicadas. Se han encontrado las siguientes:

SC.5.1. *Comparar todas las distribuciones de frecuencias relativas condicionales de una variable para los distintos valores de la otra variable.* El estudiante utiliza implícitamente la siguiente propiedad: La dependencia de una variable *B* respecto a otra variable *A* lleva a la variación de las frecuencias relativas condicionales $h(B/A)$ cuando *A* cambia. Hemos considerado en esta estrategia los estudiantes que comprueban la igualdad: $a/(a+b)=c/(c+d)$ o bien utilizan un procedimiento semejante, comparando distribuciones condicionales por columnas. Jenkins y Ward (1965) fueron los que propusieron esta estrategia como posible solución correcta, que también aparece en Estepa (1993). En el siguiente ejemplo (ver análisis en la Tabla 3.10.3.23), donde el estudiante indica que *no hay* asociación, se comienza ampliando la tabla mediante el cálculo de las frecuencias marginales por filas. Posteriormente el estudiante calcula el porcentaje de personas que padecen estrés entre los sujetos que no tienen insomnio. A continuación, plantea una regla de tres donde relaciona las frecuencias absolutas con las frecuencias marginales para calcular el número de sujetos que no tienen insomnio y tienen estrés. Al obtener el mismo valor, llega a la conclusión de que hay el mismo porcentaje de personas con estrés entre los sujetos con y sin insomnio.

	<i>Estrés</i>	<i>No estrés</i>	<i>Total</i>
<i>No tiene insomnio</i>	60	40	100
<i>Tiene insomnio</i>	90	60	150

“Si observamos esta tabla se ve claramente que el 60% de las personas sin insomnio padece estrés. Resultando una regla de tres:

$$\begin{array}{l}
 90 \rightarrow 150 \\
 ? \rightarrow 100
 \end{array}
 \quad
 \frac{90 \times 100}{150} = 60$$

Se observa que 90 es el 60% de 150 es decir hay el mismo porcentaje de personas con estrés, con o sin insomnio” (Estudiante 28, ítem 1).

Tabla 3.10.3.23. Análisis semiótico de la estrategia SC.5.1.

U	Expresión			Contenido	
3		<i>estrés</i>	<i>no estrés</i>	<i>total</i>	<ul style="list-style-type: none"> • Calcula y añade a la tabla las frecuencias marginales por fila (concepto) y procedimiento. Particularización a este caso y representación de resultados.
	<i>No tiene insomnio</i>	60	40	100	
	<i>tiene insomnio</i>	90	60	150	
4	<p>“Si observamos esta tabla se ve claramente que el 60% de las personas sin insomnio padece estrés. Resultando una regla de tres: $90 \rightarrow 150$ $\frac{90 \times 100}{150} = 60$ $? \rightarrow 100$</p> <p>Se observa que 90 es el 60% de 150 es decir hay el mismo porcentaje de personas con estrés, con o sin insomnio”.</p>				<ul style="list-style-type: none"> • Calcula el porcentaje de personas sin insomnio de los que padecen estrés, es decir, expresa en porcentaje la frecuencia relativa condicional (procedimiento y conceptos). • Aplica una regla de tres (procedimiento); puesto que los términos guardan proporción (propiedad). • Usa los símbolos de multiplicar, dividir (operaciones) e igualdad, flechas para representar una relación (propiedad) y una interrogación que representa una incógnita (representación). • Repite los cálculos, aunque no los escribe para las personas con insomnio. • Compara los porcentajes de personas con estrés con y sin insomnio (en dos distribuciones condicionales; procedimiento). • Como son iguales, deduce la independencia (aplica una propiedad para justificar un concepto, particularizando a este caso). • Aplica un argumento deductivo

SC.5.2. Comparación de posibilidades a favor y en contra de B en cada valor de

A. En esta estrategia, descrita por Estepa (1993), el estudiante emplearía implícitamente la propiedad de que existe una correspondencia unívoca entre la probabilidad de un suceso y la razón de sus probabilidades a favor y en contra. Hemos considerado en esta estrategia los estudiantes que comprueban la igualdad: $(a/c)=(b/d)$ o bien aplican un método equivalente por columnas. En el ejemplo, que se incluye a continuación (analizado en la Tabla 3.10.3.24) y donde el estudiante considera que hay relación, se realiza una separación en dos bloques, por un lado la columna de los niños problemáticos y por otro lado la de los niños sin problemas, incorporando en cada bloque el total de la columna. A continuación el estudiante establece una relación entre los valores de cada columna mediante comparación de posibilidades. Para clarificar los resultados, simplifica los valores dividiéndolos todos por 10. Finalmente se comparan los resultados de cada bloque para llegar a la conclusión obtenida.

“En el primer caso operamos sobre 140 niños, 40 con problemas y 100 sin ellos, simplificando podemos decir que 4 de cada 14 niños con hermanos presentarán problemas (hemos dividido todas las cifras entre 10 para simplificar). Partiendo de lo hecho arriba podemos afirmar que 10 de cada 11 hijos únicos presentan problemas” (Estudiante 21, ítem 2).

Tabla 3.10.3.24. Análisis semiótico de la estrategia SC.5.2.

U	Expresión	Contenido
3	<p>“En el primer caso operamos sobre 140 niños, 40 con problemas y 100 sin ellos, simplificando podemos decir que 4 de cada 14 niños con hermanos presentarán problemas (hemos dividido todas las cifras entre 10 para simplificar). Partiendo de lo hecho arriba podemos afirmar que 10 de cada 11 hijos únicos presentan problemas”.</p>	<ul style="list-style-type: none"> • Calcula el total de niños con hermanos (procedimiento y concepto; particularización y representación de resultados). • Identifica el número de casos favorables (niños con problemas) y desfavorables (sin ellos) entre los casos posibles (niños con hermanos). Restringe el espacio muestral) (procedimiento y concepto, particularizando a este caso y representando sus cálculos). • Reconoce el hecho de que dividir todas las celdas por la misma cifra no afecta para reconocer la asociación (propiedad; particulariza la proporcionalidad). Divide todas las celdas por 10 para simplificar (procedimiento). • Calcula las razones de niños con problemas, entre los niños con y sin hermanos (procedimiento). • Compara estas razones; como hay diferencias, deduce la dependencia (aplica una propiedad para justificar un concepto). • Aplica un argumento deductivo.

SC.5.3. Calcular un coeficiente para medidas de asociación en tablas 2x2.

Hemos encontrado un estudiante que calcula la razón de productos cruzados, que es un coeficiente de asociación para tablas 2x2 (Ato y López, 1996; Batanero y Díaz, 2008). Posiblemente se trate de un estudiante repetidor que ha estudiado el tema en el curso anterior. No la hemos encontrado descrita en los antecedentes. En el ejemplo que se analiza en la Tabla 3.10.3.25 el estudiante considera que no hay relación y comienza haciendo una asignación de letras a los valores de las celdas (a , b , c y d). Para saber si existe asociación entre las variables calcula el valor de la razón de productos cruzados, obteniendo el valor 1. Finalmente concluye por este valor que no existe relación.

“Para saber si es correcto que tener insomnio es un factor desencadenante del estrés. Calculando la OR podremos saber si es un factor o no.

$$OR = \frac{AxD}{BxC} = \frac{90x40}{60x60} = \frac{3600}{3600} = 1$$

No existe relación entre causa y efecto. Sin significación.” (Estudiante 221, ítem 1).

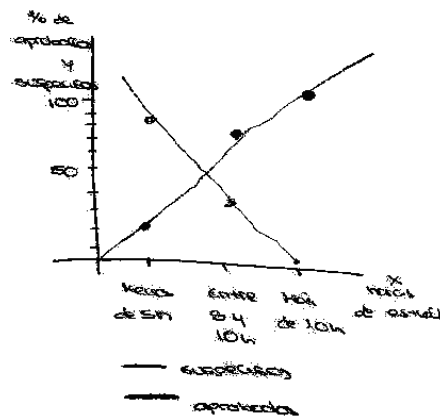
Tabla 3.10.3.25. Análisis semiótico de la estrategia SC.5.3.

U	Expresión	Contenido
3	<p>“Para saber si es correcto que tener insomnio es un factor desencadenante del estrés. Calculando la OR podremos saber si es un factor o no.”</p>	<ul style="list-style-type: none"> • Identifica el problema como de estudio de la asociación y recuerda uno de los coeficientes de asociación. Utiliza el símbolo adecuado OR (conceptos, y procedimiento, particularización).
4	<p>“$OR = \frac{AxD}{BxC} = \frac{90x40}{60x60} = \frac{3600}{3600} = 1$”</p>	<ul style="list-style-type: none"> • Utiliza símbolos para representar cada celda: A, B, C y D (particularización, concepto y representación). • Utiliza la fórmula de “razón de productos cruzados, que implica multiplicaciones y divisiones (procedimientos). • Calcula y compara dos razones (concepto y procedimiento).
5	<p>“No existe relación entre causa y efecto. Sin significación.”</p>	<ul style="list-style-type: none"> • Recuerda que un valor 1 de este coeficiente significa que no existe relación (particularización de una propiedad). • Aplica un argumento deductivo.

Estrategias parcialmente correctas

SP.5.4. El estudiante representa gráficamente las distribuciones condicionales por filas, estudiando el crecimiento/ decrecimiento de las frecuencias en cada distribución condicional. En el siguiente ejemplo (analizado en la Tabla 3.10.3.26), donde el estudiante indica que hay relación, se representa correctamente un gráfico bivalente, cuya complejidad semiótica es alta según Arteaga (2011). El estudiante utiliza frecuencias absolutas dobles y marginales, así como frecuencias relativas condicionales. Reconoce la importancia de comparar las distribuciones condicionales, para estudiar la asociación, y analiza el crecimiento/decrecimiento de estas frecuencias. No hemos encontrado esta estrategia en los trabajos previos. Aunque en el ítem dado funciona bien, porque efectivamente hay una relación lineal entre tiempo de estudio y número de aprobados o suspensos, la consideramos parcialmente correcta, porque no funcionaría en el caso general

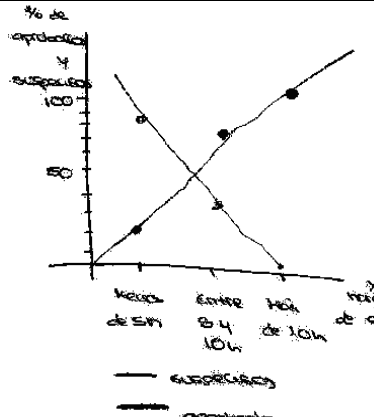
“menos de 5 h
 $5/25=0,2$ 20% aprueban
 $20/25=0,8$ 80% suspenden
entre 5 y 10 h
 $15/22=0,68$ 68% aprueban
 $7/22=0,31$ 31% suspenden
Más de 10 h
 $51/53=0,96$ 96% aprueban
 $2/53=0,03$ 3% suspenden



En la gráfica se observa claramente como a medida que aumenta las horas de estudio aumenta al número de aprobados” (Estudiante 30, ítem 4).

Tabla 3.10.3.26. Análisis semiótico de la estrategia SP.5.4.

U	Expresión	Contenido
3	<p><u>“menos de 5 h</u> $5/25=0,2$ 20% aprueban $20/25=0,8$ 80% suspenden <u>entre 5 y 10 h</u> $15/22=0,68$ 68% aprueban $7/22=0,31$ 31% suspenden <u>Más de 10 h</u> $51/53=0,96$ 96% aprueban $2/53=0,03$ 3% suspenden”</p>	<ul style="list-style-type: none"> • Identifica las tres distribuciones condicionales de la variable tiempo de estudio (concepto; particularización). • Identifica la frecuencia marginal en cada distribución (concepto; particularización). • Calcula el porcentaje de aprobados y suspensos en cada una de las tres distribuciones (procedimiento particularización). • Utiliza y representa divisiones, signos de igualdad y % (conceptos; representación).
4		<ul style="list-style-type: none"> • Representa un gráfico donde aparece en el eje x, la variable tiempo de estudio, y en el eje y los porcentajes de aprobados y suspensos calculados anteriormente (representación, procedimiento). • A los tres momentos del eje x le hace corresponder dos puntos (% de aprobados y % de suspensos).

		<ul style="list-style-type: none"> • Une mediante una línea los porcentajes de aprobados, y mediante otra línea los porcentajes de suspensos (interpolación, concepto y procedimiento).
5	<p>“En la gráfica se observa claramente como a medida que aumenta las horas de estudio aumenta el nº de aprobados”.</p>	<ul style="list-style-type: none"> • Observa el crecimiento y decrecimiento de la gráfica (concepto y aplicación de una propiedad). • Se fija en la línea que tiene un crecimiento continuo (particulariza una propiedad). • Aplica un argumento deductivo.

SP.5.5. *Supone que, para que se dé la independencia todas las frecuencias relativas dobles de la tabla deben ser iguales* (es decir el 25% de casos en cada celda). No hemos encontrado descrita esta estrategia en los antecedentes. Un ejemplo se presenta a continuación y se analiza en la Tabla 3.10.3.27. El estudiante, que piensa que hay asociación, utiliza frecuencias absolutas dobles, y divide el total por el número de celdas, suponiendo incorrectamente que las frecuencias en todas las celdas serían iguales en el caso de independencia. Compara estas frecuencias “esperadas” con las observadas y al ser diferentes deduce la asociación. Consideramos que la estrategia es parcialmente correcta al incluir la comparación entre frecuencias observadas y esperadas, pero el modo de cálculo de las esperadas solo sería válido si las frecuencias marginales de las filas y las columnas fuesen iguales.

“He dividido 250 entre 4 (25%) para ver el número de personas que tendría que haber en cada cuadrado (25%) si no afectara tener insomnio. Tras esto he visto que las personas que no tienen insomnio y padece estrés está cerca del 25%, igual que las personas que tienen insomnio y no padece estrés. Las personas que no tienen insomnio ni estrés son menos del 25% y los que tienen insomnio y estrés sobrepasan mucho el 25%” (Estudiante 1, ítem 1).

Tabla 3.10.3.27. Análisis semiótico de la estrategia SP.5.5.

U	Expresión	Contenido
3	<p>“He dividido 250 entre 4 (25%) para ver el número de personas que tendría que haber en cada cuadrado (25%) si no afectara tener insomnio. Tras esto he visto que las personas que no tienen insomnio y padece estrés está cerca del 25%, igual que las personas que tienen insomnio y no padece estrés. Las personas que no tienen insomnio ni estrés son menos del 25% y los que tienen insomnio y estrés sobrepasan mucho el 25%”</p>	<ul style="list-style-type: none"> • Divide el total de la muestra 250 entre el número de celdas (procedimiento). <i>Conflicto</i> al considerar que deberían ser idénticas en caso de independencia • Compara en cada celda entre frecuencias observadas y “esperadas” (procedimiento). • Identifica las celdas que están próximas o son diferentes al 25% (diferencias entre frecuencias observadas y esperadas)

<p><i>no padece estrés. Las personas que no tienen insomnio ni estrés son menos del 25% y los que tienen insomnio y estrés sobrepasan mucho el 25%”.</i></p>	<ul style="list-style-type: none"> • Como hay diferencia, atribuye asociación (particulariza correctamente una propiedad). El <i>conflicto</i> se produce porque opera como si las frecuencias marginales fuese iguales. Representa mediante símbolos los porcentajes (representación de conceptos y procedimientos). • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas.
--	---

Estrategias incorrectas

SI.5.6. *Comparar las frecuencias dobles relativas entre si o con el número total de observaciones.* En este caso el estudiante, primeramente calcula todas las frecuencias relativas y luego las compara entre sí. En el ejemplo reproducido a continuación de forma particular elige las celdas de mayor frecuencia, y al ser mayor que el resto, considera que *si hay relación*, estrategia descrita por Estepa (1993) (Ver análisis en la Tabla 3.10.3.29). El conflicto se produce al no diferenciar entre frecuencias dobles y condicionales y sería similar al descrito por Falk (1986) en relación a la probabilidad condicional, donde los estudiantes confunden probabilidad condicional y conjunta.

“He usado los datos de cada variable, y he hecho el % de cada uno de ellos respecto al total; y he comparado los resultados

$$\frac{90}{250}100 = 36\% \quad \frac{60}{250}100 = 24\%$$

$$\frac{60}{250}100 = 24\% \quad \frac{40}{250}100 = 16\%$$

Tiene insomnio y padece estrés” (Estudiante 11, ítem 1).

Tabla 3.10.3.29. Análisis semiótico de la estrategia SI.5.6.

U	Expresión	Contenido
3	<p><i>“He usado los datos de cada variable, y he hecho el % de cada uno de ellos respecto al total; y he comparado los resultados”.</i></p>	<ul style="list-style-type: none"> • Calcula el total de sujetos que participan en el experimento, las frecuencias relativas dobles o porcentaje de sujetos presentes en cada celda (procedimientos y conceptos). • Utiliza símbolo de porcentaje % (concepto y representación). • Compara los resultados, buscando el porcentaje mayor de ellos (procedimiento).
4	<p><i>“ $\frac{90}{250}100 = 36\%$ $\frac{60}{250}100 = 24\%$ $\frac{60}{250}100 = 24\%$ $\frac{40}{250}100 = 16\%$ Tiene insomnio y padece estrés”</i></p>	<ul style="list-style-type: none"> • Utiliza multiplicaciones y divisiones (conceptos y procedimientos). • Los escribe en la misma posición en la que están en la tabla como método de identificación (procedimiento). • Utiliza un círculo para resaltar el valor superior de los porcentajes (representación). • Asume incorrectamente que para que se de la independencia los porcentajes debieran ser iguales (<i>conflicto</i> al aplicar una propiedad). • Otro <i>conflicto</i> es no diferenciar entre frecuencias dobles y condicionales • Aplica un argumento deductivo

SI.5.7. *Calcula las frecuencias relativas marginales y las compara con alguna condicional, cometiendo un error.* Es una estrategia incorrecta, debido al error, como vemos en el ejemplo que se presenta a continuación y se analiza en la Tabla 3.10.3.30 y donde el estudiante indica que *si hay relación*. 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). Debido a este conflicto, llega a una solución incorrecta.

“A=ser hijo único $\rightarrow P(A)=0,44$

B=ser un niño problemático $\rightarrow P(B)=0,56$, $n=250$, $P(B)=P(A/B)=>0,56=0,785$ ”

La relación entre tener hermanos y ser problemáticos debe ser mayor que el anterior ejercicio ya que la diferencia de datos entre sucesos no A, A y B, no B es mayor” (Estudiante 414, ítem 2).

Tabla 3.10.3.30. Análisis semiótico de la estrategia SI.5.7.

U	Expresión	Contenido
3	<i>“A=ser hijo único $\rightarrow P(A)=0,44$ B=ser un niño problemático $\rightarrow P(B)=0,56$”</i>	<ul style="list-style-type: none"> • El estudiante utiliza símbolos: A, B, flechas, $P(A)$, $P(B)$ y signo de igualdad (conceptos y representaciones). • El estudiante calcula las probabilidades simples, colocando una debajo de la otra para facilitar su comparación (concepto y procedimiento).
4	<i>“$n=250$ $P(B)=P(A/B)=>0,56=0,785$”</i>	<ul style="list-style-type: none"> • El estudiante identifica el total de la muestra (concepto). • El estudiante intenta aplicar la propiedad de independencia de sucesos de $P(A)=P(A/B)$ (propiedad). • El estudiante confunde $P(B/A)$ con $P(A/B)$ (conflicto).
5	<i>“La relación entre tener hermanos y ser problemáticos debe ser mayor que el anterior ejercicio ya que la diferencia de datos entre sucesos no A, A y B, no B es mayor.”</i>	<ul style="list-style-type: none"> • Compara los valores de no A con A, y los valores de B con no B, que son los valores que asigno a las frecuencias relativas marginales (procedimiento). • Aplica un argumento deductivo que contiene un proceso de síntesis de los cálculos realizados y relaciones observadas. • Realiza varios procesos de representación y particularización.

SI.5.8. *Otros procedimientos.* En casos esporádicos se usan procedimientos que no se especifica con claridad, debido a la dificultad de expresión del estudiante, aunque se ve claro el uso de todos los datos, a nivel 5. Un ejemplo se presenta a continuación y se analiza en la Tabla 3.10.3.31. El estudiante considera en este ejemplo que si hay asociación. Junto con propiedades de la asociación correctamente identificadas y aplicadas, los estudiantes que usan esta estrategia aplican propiedades incorrectas. En general, el principal problema es usar las frecuencias absolutas en lugar de las relativas para comparar dos distribuciones. Esta estrategia sería correcta si los tamaños de muestra fuesen iguales, pero no lo son al ser diferentes. Según Konold, Pollatsek, Well y Gagnon (1997) esta estrategia indica que el estudiante aún no ha adquirido una comprensión suficiente del concepto de distribución. No hemos encontrado esta estrategia en los antecedentes.

TABLA DE FRECUENCIA (APROBADOS)

X	F	p_i	F	P_i
-5h	5	0,07	5	7
5-10h	15	0,21	20	21
+10h	<u>51</u>	<u>0,72</u>	71	<u>72</u>
	$N=71$	1		100%

TABLA DE FRECUENCIA (SUSPENSOS)

X	F	p_i	f	P_i
-5h	20	0,68	20	68
5-10h	7	0,24	27	24
+10h	<u>2</u>	<u>0,07</u>	29	<u>7</u>
	$N=25$	1		99%

(Estudiante 19, ítem 4).

Tabla 3.10.3.31. Análisis semiótico de la estrategia SI.5.8.

U	Expresión	Contenido																									
3	<p>“TABLA DE FRECUENCIA (APROBADOS)”</p> <table border="1"> <thead> <tr> <th>X</th> <th>F</th> <th>p_i</th> <th>F</th> <th>P_i</th> </tr> </thead> <tbody> <tr> <td>-5h</td> <td>5</td> <td>0,07</td> <td>5</td> <td>7</td> </tr> <tr> <td>5-10h</td> <td>15</td> <td>0,21</td> <td>20</td> <td>21</td> </tr> <tr> <td>+10h</td> <td><u>51</u></td> <td><u>0,72</u></td> <td>71</td> <td><u>72</u></td> </tr> <tr> <td></td> <td>$N=71$</td> <td>1</td> <td></td> <td>100%</td> </tr> </tbody> </table>	X	F	p_i	F	P_i	-5h	5	0,07	5	7	5-10h	15	0,21	20	21	+10h	<u>51</u>	<u>0,72</u>	71	<u>72</u>		$N=71$	1		100%	<ul style="list-style-type: none"> • Estudia la distribución de los aprobados (procedimiento). • Transforma la tabla de contingencia en una tabla de frecuencia (procedimiento). • Utiliza los símbolos x, f_i, p_i, f_a y P_i para representar lo que es cada columna (propiedad y conceptos). Proceso de representación. • Calcula las probabilidades de cada celda, comprobando que suman uno (Particularización de procedimiento y concepto). • Calcula las frecuencias absolutas acumuladas (Particularización de procedimiento y concepto). • Calcula las probabilidades en porcentaje, comprobando que suman cien (Particularización de propiedades, procedimiento y concepto). • Utiliza el símbolo % (operaciones). Representación
X	F	p_i	F	P_i																							
-5h	5	0,07	5	7																							
5-10h	15	0,21	20	21																							
+10h	<u>51</u>	<u>0,72</u>	71	<u>72</u>																							
	$N=71$	1		100%																							
4	<p>“TABLA DE FRECUENCIA (SUSPENSOS)”</p> <table border="1"> <thead> <tr> <th>X</th> <th>F</th> <th>p_i</th> <th>f</th> <th>P_i</th> </tr> </thead> <tbody> <tr> <td>-5h</td> <td>20</td> <td>0,68</td> <td>20</td> <td>68</td> </tr> <tr> <td>5-10</td> <td>7</td> <td>0,24</td> <td>27</td> <td>24</td> </tr> <tr> <td>+10h</td> <td><u>2</u></td> <td><u>0,07</u></td> <td>29</td> <td><u>7</u></td> </tr> <tr> <td></td> <td>$N=25$</td> <td>1</td> <td></td> <td>99%</td> </tr> </tbody> </table>	X	F	p_i	f	P_i	-5h	20	0,68	20	68	5-10	7	0,24	27	24	+10h	<u>2</u>	<u>0,07</u>	29	<u>7</u>		$N=25$	1		99%	<ul style="list-style-type: none"> • Repite el procedimiento anterior para la distribución de los suspensos (procedimiento).
X	F	p_i	f	P_i																							
-5h	20	0,68	20	68																							
5-10	7	0,24	27	24																							
+10h	<u>2</u>	<u>0,07</u>	29	<u>7</u>																							
	$N=25$	1		99%																							

Resumen de configuraciones cognitivas

Una vez categorizadas las estrategias, se analizó la frecuencia con que se presenta cada una en los diferentes ítems (Tabla 3.10.3.32). Se observa una gran variedad de estrategias; algunas de ellas muy idiosincráticas, al haber sido utilizada por muy pocos estudiantes.

Entre las estrategias incorrectas más frecuentes encontramos el uso único de la celda de mayor frecuencia (SI.1.1 (celda a) y SI.1.2 (celda diferente de la a)), para el nivel 1 estrategia que fue descrita por Inhelder y Piaget (1955), Smedlund (1963), Beyth y Marom (1982), Shaklee y Mins (1982), Yates y Curley (1986), y encontrada en Estepa (1993).

Tabla 3.10.3.32. Frecuencia (y porcentaje) de estudiantes según la estrategia

		Tablas 2x2			Tabla 2x3	Total	
		Ítem 1	Ítem 2	Ítem 3	Ítem 4		
Nivel 0 Estrategias incorrectas	SI.0.1.	6 (1,45)	7 (1,7)	5 (1,2)	4 (1)	22 (1,3)	
	SI.0.2.	7 (1,69)	8 (1,9)	10 (2,4)	4 (1)	29 (1,8)	
Nivel 1 Estrategias incorrectas	SI.1.1.	69 (16,67)	0 (0)	33 (8)	0 (0)	102 (6,2)	
	SI.1.2.	-	20 (4,8)	-	7 (1,7)	27 (1,6)	
	SI.1.3.	1 (0,24)	0 (0)	0 (0)	1 (0,2)	2 (0,1)	
	SI.1.4.	2 (0,48)	0 (0)	0 (0)	0 (0)	2 (0,1)	
	SI.1.5.	1 (0,24)	0 (0)	0 (0)	0 (0)	1 (0,0)	
Nivel 2 Estrategia p. correcta	SI.2.1.1				93 (22,5)	93 (5,6)	
Estrategias incorrectas	SI.2.1.1.	80 (19,32)	68 (16,4)	68 (16,4)		216 (13,0)	
	SI.2.1.2.	1 (0,24)	25 (6,0)	5 (1,2)		31 (1,9)	
	SI.2.2.	24 (5,8)	61 (14,7)	79 (19,1)	8 (1,9)	172 (10,4)	
	SI.2.3.	3 (0,72)	0 (0)	1 (0,2)	14 (3,4)	18 (1,1)	
	SI.2.4.	0 (0)	0 (0)	0 (0)	4 (1)	4 (0,2)	
Nivel 3 Estrategias p. correctas	SP.3.1.	-	-	-	48 (11,6)	48 (2,9)	
	Estrategias incorrectas	SI.3.2.	23 (5,56)	6 (1,5)	4 (1)	-	33 (2,0)
		SI.3.3.	4 (0,97)	6 (1,5)	5 (1,2)	-	15 (0,9)
		SI.3.4.	0 (0)	4 (1)	0 (0)	-	4 (0,2)
		SI.3.5.	-	-	-	1 (0,2)	1 (0,0)
Nivel 4 Estrategias p. correctas	SP.4.1.	0 (0)	5 (1,2)	4 (1)	-	9 (0,5)	
	SP.4.2.	73 (17,63)	107 (25,9)	94 (22,7)	153 (37)	427 (25,8)	
	SP.4.3.	3 (0,72)	3 (0,7)	2 (0,5)	1 (0,2)	9 (0,5)	
	Estrategias incorrectas	SI.4.4.	1 (0,24)	0 (0)	1 (0,2)	-	2 (0,1)
		SI.4.5.	17 (4,11)	3 (0,7)	8 (1,9)	2 (0,5)	30 (1,8)
		SI.4.6.	6 (1,45)	4 (1)	6 (1,5)	3 (0,7)	19 (1,1)
		SI.4.7.	3 (0,72)	5 (1,2)	4 (1)	0 (0)	12 (0,7)
Nivel 5 Estrategias correcta	SC.5.1.	37 (8,94)	32 (7,7)	26 (6,3)	24 (5,8)	119 (7,2)	
	SC.5.2.	8 (1,93)	5 (1,2)	1 (0,2)	3 (0,7)	17 (1,0)	
	SC.5.3.	1 (0,24)	0 (0)	1 (0,2)	0 (0)	2 (0,1)	
	Estrategias p. correctas	SP.5.4.	9 (2,17)	8 (1,9)	6 (1,5)	7 (1,7)	30 (1,8)
		SP.5.5.	1 (0,24)	0 (0)	0 (0)	0 (0)	1 (0,0)
	Estrategias incorrectas	SI.5.6.	16 (3,86)	24 (5,8)	19 (5,6)	12 (2,9)	71 (4,3)
		SI.5.7.	1 (0,24)	2 (0,5)	2 (0,5)	1 (0,2)	6 (0,4)
		SI.5.8.	3 (0,72)	3 (0,7)	4 (1)	2 (0,5)	12 (0,7)
No responde		15 (3,62)	8 (1,9)	25 (6,0)	22 (5,3)	70 (4,2)	
Total		414	414	414	414	1656	

Para el nivel 2 fue el uso de una distribución condicional (SI.2.1.1), consistente en comparar la celda a (presencia de los dos caracteres) con las celdas b y c (presencia de uno de los dos caracteres). Recordemos que el uso de esta estrategia también ha sido frecuente en investigaciones previas como, por ejemplo, Smedlund (1963), Beyth y Marom (1982), Shaklee y Mins (1982) o Yates y Curley (1986). También destaca en el nivel 2 la estrategia SI.2.2, consistente en comparar las frecuencias en una de las diagonales de la tabla.

Entre las parcialmente correctas, las más frecuentes son para el nivel 3, la estrategia SP.3.1 consistente en comparar dos distribuciones condicionales por columnas, ignorando una distribución condicional en una tabla 2x3. Estepa (1993) encuentra un estudiante con esta estrategia. Para el nivel 4, tiene mayor frecuencia la

estrategia SP.4.2, comparar dos distribuciones condicionales en forma aditiva, que es una de las estrategias encontradas en el estudio de Inhelder y Piaget (1955), Estepa (1993) o Godino y Batanero (1998). Por último, para el nivel 5, destaca la estrategia SP.5.4 donde el estudiante representa gráficamente las distribuciones condicionales por filas.

Entre las correctas, para el nivel 5, la estrategia SC.5.1 de comparar las distribuciones de frecuencias relativas condicionales de una variable, es la que aparece más. Jenkins y Ward (1965) propusieron esta estrategia como posible solución correcta, y se describe también en la investigación de Estepa (1993).

En la investigación de Estepa las estrategias más frecuentes fueron: utilizar una única distribución condicional (entre las estrategias incorrectas); comparar posibilidades o razón (entre las estrategias correctas); y comparar una frecuencia relativa de cada distribución (entre las estrategias parcialmente correctas). La más frecuente entre estas resulta ser la estrategia incorrecta, seguida por la estrategia parcialmente correcta y por último la estrategia correcta.

Una síntesis de la anterior, es la Tabla 3.10.3.33, donde se resumen las frecuencias de cada una de los distintos tipos de estrategias (incorrectas, parcialmente correctas y correctas) para cada uno de los niveles. Para los niveles 0 y 1 todas las estrategias utilizadas son incorrectas por lo que de nuevo se confirma que estas estrategias no son productivas. Lo mismo ocurre en el nivel 2 donde a lo más son parcialmente correctas. En todas ellas se manifiesta la concepción local de la asociación, que deberá ser superada por los estudiantes. Observamos que, aunque no hay una diferencia apreciable de uso del nivel 0 en los cuatro ítems, la suma de los niveles 1 y 2, así como el nivel 1 es mucho más frecuente en el ítem 1, donde las teorías previas de los estudiantes contradicen la asociación de los datos.

En los niveles 3 y 4, aparecen estrategias parcialmente correctas e incorrectas; en el nivel 3 la mayoría son incorrectas; tan sólo en el último ítem aparece un porcentaje de casos en que en este nivel los estudiantes dieron estrategias parcialmente correctas, ya que en 48 casos se comparan dos distribuciones condicionales por columnas, ignorando una distribución condicional en tablas 2x3. Por último señalamos que sólo en el nivel 5, aparecen estrategias de los tres tipos mencionados y en particular, estrategias correctas, que son mayoría (a pesar de su escasa frecuencia) en este nivel. Es por tanto necesario llevar al estudiante a usar todos los datos y además, relacionarlos con estrategias multiplicativas.

Tabla 3.10.3.33. Frecuencia (y porcentaje) de estudiantes según el tipo de estrategia

		Ítem 1	Ítem 2	Ítem 3	Ítem 4	Total ítems
Nivel 0	Incorrectas	13 (3,1)	15 (3,6)	15 (3,6)	8 (1,9)	51 (3,1)
Nivel 1	Incorrectas	73 (17,6)	20 (4,8)	33 (8,1)	8 (1,9)	134 (8,1)
Nivel 2	Incorrectas	108 (25,8)	154 (37,2)	153 (37)	26 (5,5)	441 (26,6)
	P. correctas				93 (22,5)	93 (5,6)
Nivel 3	Incorrecta	27 (6,6)	16 (3,8)	9 (2,2)	1 (0,2)	53 (3,2)
	P. correcta	0 (0)	0 (0)	0 (0)	48 (11,6)	48 (2,9)
Nivel 4	Incorrectas	27 (6,6)	12 (2,9)	19 (4,5)	5 (1,2)	63 (3,8)
	P. correcta	76 (18,3)	115 (27,8)	100 (24,3)	151 (37,2)	445 (26,9)
Nivel 5	Incorrectas	20 (4,8)	29 (7,1)	25 (6,1)	15 (3,5)	89 (5,4)
	P. correcta	10 (2,4)	8 (1,9)	6 (1,4)	7 (1,7)	31 (1,9)
	Correcta	46 (11,1)	37 (9)	28 (6,8)	27 (6,6)	138 (8,3)
No responde		15 (3,6)	8 (1,9)	25 (6)	22 (5,3)	70 (4,2)
Total		414	414	414	414	1656

Estepa (1993) indica que en su investigación un número importante de estudiantes trabajaron en el nivel 1 en tablas 2x2, pero pasan a un nivel superior al aumentar la dimensión de la tabla, lo que nos ocurre en nuestro caso ya que en el nivel 1 la frecuencia más baja nos aparece en el ítem 4. Estepa informa que para el nivel 2 se duplican los resultados en tablas rxc respecto a las tablas 2x2, aunque en nuestro caso no, ya que nos ocurre exactamente lo contrario. El nivel 3 fue muy limitado en ambos tipos de tablas, y también poco frecuente en nuestro caso, donde la máxima frecuencia aparece en el ítem 4 de tabla rxc. El nivel 4 fue para Estepa similar en ambos tipos de tablas, y para nosotros más frecuente en tablas rxc. Finalmente, para el nivel 5, Estepa informa de la disminución de la frecuencia en tablas rxc, al igual que en nuestro caso.

En la Tabla 3.10.3.34 se presenta un nuevo resumen de las estrategias; en este caso tan sólo tenemos en cuenta su corrección. Se observa una abrumadora mayoría de estrategias incorrectas en todos los ítems, excepto el 4 y en su conjunto. Destaca, aunque no con mucha diferencia la que corresponde con el ítem 1 que, como hemos dicho contradice las teorías previas.

Tabla 3.10.3.34. Frecuencia (y porcentaje) de estudiantes según el tipo de estrategia

	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Total ítems
Incorrectas	267 (64,6)	246 (59,4)	255 (61,5)	63 (15,2)	924 (55,9)
P. correctas	86 (20,7)	123 (29,7)	106 (25,7)	295 (71,2)	524 (31,6)
Correctas	46 (11,1)	37 (9)	28 (6,8)	27 (6,6)	138 (8,3)
No responde	15 (3,6)	8 (1,9)	25 (6)	22 (5,3)	70 (4,2)

Es también muy alta (entre el 20 y 70%) la frecuencia de estrategias parcialmente correctas, especialmente en el último ítem, donde hubo menos incorrectas. Son muy pocas las estrategias correctas; siendo el caso de independencia donde los resultados son más altos. La frecuencia de no respuesta es muy pequeña; deducimos con ello el interés de los estudiantes en responder a los problemas propuestos.

Una conclusión es la escasa competencia para realizar juicios de asociación con que abordan los estudiantes sus estudios de psicología; dicha incompetencia les dificultará la realización de estos juicios en su vida profesional. La experiencia de enseñanza que se describe y evalúa en los siguientes capítulos está precisamente orientada a preparar a los estudiantes en el análisis de las tablas de contingencia y la emisión de juicios de asociación adecuados.

Por otro lado, aunque el estudio de las frecuencias con que se producen las diferentes estrategias ya nos proporciona una información de interés no nos da una explicación del por qué las estrategias correctas no se han desarrollado en forma intuitiva en los estudiantes, sin necesidad de instrucción previa, como se daba por supuesto en los estudios de Inhelder y Piaget.

Una posible explicación nos la proporciona el análisis de los objetos matemáticos implícitos en cada una de las estrategias, que presentamos en las Tablas 3.10.3.35 a 3.10.3.38. No se han incluido las estrategias SI.0.1 y SI.0.2, ya que no utilizan los datos de la tabla. Tampoco se han incluido las estrategias SI.1.5, SI.2.4, SI.3.5, SI.4.7 y SI.5.8, ya que son las estrategias que corresponden a “otros procedimientos”. Cada estrategia usa diferentes objetos matemáticos, por tanto, en términos de Godino (2002b) se trataría de configuraciones cognitivas diferenciadas. Pero la cantidad y dificultad de objetos implicados no es la misma en todas ellas.

Tabla 3.10.3.35. Conceptos implícitos en las diferentes estrategias

Conceptos	Nivel 1				Nivel 2				Nivel 3				Nivel 4						Nivel 5							
	SI	SI	SI	SI	SI	SI	SI	SI	SP	SI	SI	SI	SP	SP	SP	SI	SI	SI	SC	SC	SC	SP	SP	SI	SI	
	1	2	3	4	1.1	1.2	2	3	1	2	3	4	1	2	3	4	5	6	1	2	3	4	5	6	7	
Frecuencias abs. Doble	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	
Frecuencias abs. Marginal													X	X				X					X	X		X
Frecuencias rel. Doble	X									X	X	X			X					X					X	X
Frecuencias rel. Marginal																			X							
Frecuencias rel. condicional																			X				X			
Distribución marginal															X											
Distribución condicional					X	X			X	X	X	X		X			X		X	X	X	X				
Casos favorables	X						X	X						X		X	X		X	X						
Casos desfavorables			X	X			X	X						X		X	X		X	X						
Total/casos posibles	X	X	X	X			X						X	X		X	X	X	X	X		X			X	X
Variable simple /valor					X	X							X									X	X	X		
Variable doble/valor	X	X	X	X									X	X	X	X	X	X	X	X	X	X	X	X	X	X
Porcentaje	X																		X			X		X	X	X
Razón																		X	X	X	X	X	X	X		
Proporción																			X	X	X	X	X			X
Probabilidad													X						X	X	X	X				
Monotonía				X				X	X								X									
Factor riesgo/riesgo relativo																					X					
OR/RR																					X					
Moda		X																								
Media															X											
Gráfico																						X				
Interpolación																						X				

Tabla 3.10.3.36. Propiedades implícitas en las diferentes estrategias

Propiedades	Nivel 1				Nivel 2				Nivel 3				Nivel 4						Nivel 5						
	SI	SI	SI	SI	SI	SI	SI	SI	SP	SI	SI	SI	SP	SP	SP	SI	SI	SI	SC	SC	SC	SP	SP	SI	SI
	1	2	3	4	1.1	1.2	2	3	1	2	3	4	1	2	3	4	5	6	1	2	3	4	5	6	7
Orden (numérico o de frecuencias)			X			X	X	X																	
Moda: valor de mayor frecuencia		X						X																	
Celdas que apoyan o no la asociación	X		X	X		X				X	X	X	X			X	X								
Asociación implica diferencia distribuciones									X					X				X							
Igualdad de f. rel. condicional en independencia																		X							
Relación entre razón de posibilidades y probabilidad																			X						
Crecimiento/ decrecimiento distribución					X				X												X				
La razón de posibilidades es 1 en caso de independencia																			X						
RR=1, en caso de independencia																				X					
En caso de independencia, la gráfica no debe mostrar crecimiento/decrecimiento																					X				
Estudio asociación implica uso de todos los datos													X	X	X	X	X	X	X	X	X	X	X	X	X
Estudio asociación implica el uso de frecuencias relativas															X			X			X	X	X	X	
Proporcionalidad																			X	X					
Axioma de la unión																									

Tabla 3.10.3.38. Lenguajes implícitos en las diferentes estrategias

Lenguajes	Nivel 1				Nivel 2				Nivel 3				Nivel 4						Nivel 5						
	SI	SI	SI	SI	SI	SI	SI	SI	SP	SI	SI	SI	SP	SP	SP	SI	SI	SI	SC	SC	SC	SP	SP	SI	SI
	1	2	3	4	1.1	1.2	2	3	1	2	3	4	1	2	3	4	5	6	1	2	3	4	5	6	7
Verbal	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Numérico	X	X	X	X			X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Simbólico	X	X					X						X			X		X		X	X	X	X	X	
Tabla																		X							
Esquema /gráfico																					X				

Tabla 3.10.3.37. Procedimientos implícitos en las diferentes estrategias

Procedimientos	Nivel 1				Nivel 2				Nivel 3				Nivel 4						Nivel 5						
	SI	SI	SI	SI	SI	SI	SI	SI	SP	SI	SI	SI	SP	SP	SP	SI	SI	SI	SC	SC	SC	SP	SP	SI	SI
	1	2	3	4	1.1	1.2	2	3	1	2	3	4	1	2	3	4	5	6	1	2	3	4	5	6	7
Lectura de tabla, identificar datos	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Regla de tres																			X						
Suma y/ó resta													X			X			X						
Multiplicar y/ó dividir	X						X								X				X	X	X	X	X	X	X
Cálculo del tanto por ciento	X																		X			X	X	X	X
Comparación		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X
Simplifica fracciones																				X					
Calcula razones																				X					
Crecimiento/decrecimiento distribución					X	X			X					X				X				X			X
Fórmulas																					X				X
Calcular el total	X																		X				X	X	X
Calcular f. relativas dobles o condicionales	X																		X					X	
Calcular medias aritméticas															X										
Gráfico de distribución condicional																						X			
Buscar menor y/ó mayor frecuencia		X	X					X																	
Interpolación																						X			
Comparar distribución marginal																		X							X
Comparar distribución condicionales					X	X			X	X	X											X			
Comparar diferencia de las diagonales																X									
Interpretar “razón de productos cruzados”																					X				

Respecto al lenguaje, no se observa mucha diferencia entre estrategias correctas y otras; generalmente se usa el lenguaje verbal y numérico, con menos frecuencia el simbólico y en pocos casos gráfico. En relación a los conceptos, el número empleado es mayor conforme avanza el nivel de las estrategias, caracterizándose las de nivel 4 y 5 por el empleo de frecuencias condicionales y distribuciones condicionales, y las de nivel 5, especialmente las correctas por la razón, proporción, porcentaje y probabilidad. Es también mucho mayor el número de propiedades y procedimientos empleados en las estrategias de nivel 4 y 5, por lo que dichas estrategias tienen mayor complejidad semiótica que las de niveles anteriores.

3.10.4. CONFLICTOS SEMIÓTICOS

Una de las razones que sugieren la importancia del análisis semiótico, según Godino (2002b), es que permite identificar los conflictos semióticos de los estudiantes, esto es, las interpretaciones realizadas por los estudiantes que no son las esperadas por el profesor. Además, el análisis semiótico, describe objetos y procesos que intervienen en las prácticas matemáticas o emergen de ellos y muestran la complejidad onto-semiótica de dichas prácticas.

Este fue precisamente uno de los objetivos del análisis semiótico realizado de las estrategias de los estudiantes, por lo que, a lo largo de este capítulo se han ido destacando los encontrados en cada una de las estrategias. En las Tablas 3.10.3.1 y 3.10.3.2 se hace un resumen de los conflictos semióticos detectados en las estrategias parcialmente correctas e incorrectas, que se analizan a continuación clasificados según se refieran a la asociación, independencia y otros conceptos y propiedades.

Propiedades incorrectas atribuidas a la asociación:

- *Confundir causalidad con asociación.* En la Sección 2.2 se hizo un estudio detallado de las relaciones entre asociación y causalidad, mostrando que no siempre van ligadas, pues, aunque la causalidad implica la asociación, no sucede lo mismo en el caso contrario (Díaz y de la Fuente, 2005b). Sin embargo, los estudiantes confunden asociación y causalidad en la estrategia SI0.1, al igual que ocurrió en el trabajo de Estepa (1993) quien denomina esta creencia concepción causal de la asociación.
- *Suponer que la asociación se deduce sólo de una parte de los datos.* Implica la existencia de la *concepción local* sobre la asociación descrita por Estepa (1993), apareciendo en muchas investigaciones (Smedlund, 1963; Beyth y Marom, 1982; Shaklee y Mins, 1982; Yates y Curley, 1986; Pérez Echeverría, 1990). Aparece en

todas las estrategias basadas en la comparación de sólo una parte de los datos, es decir en todas las de nivel inferior a 4.

Tabla 3.10.3.1. Conflictos implícitos en las diferentes estrategias (niveles 0 a 3)

Tipo	Conflictos	N. 0		Nivel 1					Nivel 2					Nivel 3					
		SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI		
		1	2	1	2	3	4	5	1.1	1.2	2	3	4	1	2	3	4	5	
Conc./propiedades	Confundir probabilidades con casos favorables																	X	
	Confundir probabilidad simple y conjunta																		X
	Pensar que se pueden comparar muestras usando frecuencias absolutas				X	X	X	X	X	X					X	X	X	X	X
Asociación	Confundir causalidad con asociación (concepción causal)	X																	
	No considerar la asociación negativa (concepción unidireccional)							X											
	No usar todos los datos del enunciado (concepción local)			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Suponer asociación si $a > d$										X								
	Suponer asociación si $a > \text{total}/2$												X						
Ind.	Suponer independencia si $b = c$														X				
Otras	Probabilidad mayor que 1																		X
	No usar los datos	X	X																
	Usar las teorías previas	X																	
	Cálculos incorrectos			X															

- Interpretar la asociación inversa como independencia, es decir, mostrar la concepción unidireccional, que fue descrita por Estepa (1993), aceptando la existencia de asociación, pero no su signo. Aparece en la estrategia SI.1.5.
- Suponer que la asociación se deduce con comparaciones aditivas. Este error, que se ha descrito también en la comparación de probabilidades, es propio de los sujetos que están comenzando el periodo de las operaciones formales pero no lo han afianzado todavía y es descrito en Piaget e Inhelder (1951). Aparece en la estrategia SP.4.2 (la que ha obtenido mayor frecuencia de uso) y SI.4.4. Uno de los actos de comprensión identificado en el estudio de Estepa (1993) es que para estudiar la asociación entre dos variables, la comparación de dos o más muestras tiene que ser hecha en términos de las frecuencias relativas. Los alumnos que usan estrategias aditivas no han llegado a comprender esta propiedad.
- Suponer que hay asociación si las frecuencias absolutas condicionales son distintas. Hay asociación si las frecuencias relativas condicionales son diferentes. Los estudiantes generalizan incorrectamente una propiedad de las frecuencias relativas a

las absolutas, mediante un proceso incorrecto de generalización, encontrándose en las investigaciones de Smedlund (1963) y Shaklee y Mins (1982). Aparece en la estrategia SI.4.5.

- *Supone que la asociación se deduce de las frecuencias marginales.* Implica una confusión entre frecuencia doble y marginal, este conflicto aparece en Estepa (1993) y sería equivalente al error encontrado por Contreras (2011) entre probabilidad simple y condicional en algunos estudiantes. Aparece en las estrategias SI.4.6 y SI.5.8.

Tabla 3.10.3.2. Conflictos implícitos en las diferentes estrategias (niveles 4 y 5)

Tipo	Conflictos	Nivel 4							Nivel 5							
		SP	SP	SP	SI	SI	SI	SI	SC	SC	SC	SP	SP	SI	SI	SI
		1	2	3	4	5	6	7	1	2	3	4	5	6	7	8
Conc./Proc.	Pensar que se pueden comparar muestras usando frecuencias absolutas	X	X		X	X	X									
	Confundir frecuencia con valor de la variable			X								X				
	Confunde variable cuantitativa y cualitativa											X				
	Calcular la media de las frecuencias			X												X
	Confunde frecuencias dobles y condicionales			X										X		
	Confunde P(A B) y P(B A)														X	
	Confundir una tabla de frecuencias con una tabla de contingencia															X
Asociación	Asociación si frecuencias absolutas condicionales son distintas				X											
	Suponer que la asociación se deduce con comparaciones aditivas	X	X	X	X	X	X	X								
	Supone que la asociación se deduce de las frecuencias marginales						X									X
	Suponer la asociación se deduce de la diferencia de las diagonales				X											
Indep.	Suponer equiprobabilidad de celdas en caso de independencia												X	X		

- *Suponer la asociación se deduce de la diferencia de las celdas en las dos diagonales en la tabla 2x2.* Piaget e Inhelder indican que una estrategia (considerada correcta por los autores) es comparar la diferencia de la suma de las diagonales, estrategia que también se describe en Allan y Jenkins (1983) y Shaklee y Tucker (1980). En nuestro estudio hemos encontrado un conflicto que supone realizar la diferencia entre las celdas en cada diagonal y compararlas, es decir comparar $(a-d)$ con $(b-c)$. Este procedimiento sería incorrecto y además implica no comprender que la celda d (ausencia de los dos caracteres) tiene el mismo peso en la detección de la

asociación que la celda a (presencia de los dos). No hemos encontrado descrito este conflicto que aparece en la estrategia SI.4.4.

- *Suponer que $a > d$ en caso de asociación.* Es un caso particular de la concepción local de la asociación descrita por Estepa (1993), donde el estudiante piensa que sólo estas dos celdas influyen en la asociación. Por otro lado, al igual que en el conflicto anterior no comprenden que la celda d (ausencia de los dos caracteres) tiene el mismo peso en la detección de la asociación que la celda a (presencia de los dos). Aparece en la estrategia SI.2.2.
- *Suponer asociación si $a > total/2$,* es decir, creer que para que se presente la asociación más de la mitad de la muestra ha de tener los dos caracteres. Además de la concepción local de la asociación descrita por Estepa (1993), el estudiante piensa que sólo la presencia de los dos caracteres influye en la asociación encontrándose en muchas investigaciones el uso único de esta celda (Smedlund, 1963; Beyth y Marom, 1982; Shaklee y Mins, 1982; Yates y Curley, 1986). Aparece en la estrategia SI.2.4.

Propiedades incorrectas atribuidas a la independencia:

- *Suponer la equiprobabilidad en las celdas dobles en caso de independencia.* Supone que en caso de independencia la frecuencia ha de ser igual en todas las celdas de la tabla. Implica una confusión entre las ideas de independencia y equiprobabilidad. Sin embargo queremos resaltar que los alumnos que muestran este conflicto también presentan una idea correcta que consiste en comparar las frecuencias esperadas con las observadas para deducir la asociación, aunque el cálculo que realizan de frecuencias esperadas es incorrecto. Aparece en las estrategias SP.5.5 y SI.5.6.
- *Suponer que $b=c$ en caso de independencia.* Es un caso particular de la concepción local de la asociación (Estepa, 1993), donde el estudiante piensa que sólo estas dos celdas influyen en la asociación. Intuitivamente se identifican las celdas que serían indicativas de una asociación inversa, pero incorrectamente se supone que han de ser idénticas. Aparece en la estrategia SI.3.2.
- *Asumir incorrectamente que si la celda d que corresponde a la ausencia de los dos caracteres tiene una frecuencia elevada, ello implica independencia.* El estudiante muestra una concepción local, al trabajar sólo con una parte de los datos. Por otro lado también manifiesta la concepción determinista de la asociación descrita por

Estepa (1993) y Estepa y Batanero (1995), pues supone que no ha de haber excepciones a la asociación, por tanto en esta celda debiera haber un gran número de casos. Esta estrategia aparece en Nisbett y Ross (1980) y en nuestro estudio corresponde a la estrategia SI. 1.4.

Confusión entre conceptos y propiedades:

- *Confusión de probabilidades con casos favorables:* Se produce al calcular o comparar probabilidades. El conflicto puede ser originado por el lenguaje, pues a veces se usa como sinónimos posibilidad y probabilidad en el lenguaje ordinario, aunque no en matemáticas. Esta confusión también apareció en el trabajo de Díaz y de la Fuente (2005a) y en Contreras (2011) y aparece en la estrategia SI.3.4.
- *Confundir los conceptos de razón y frecuencia absoluta.* Relacionado con el anterior, pues la frecuencia correspondería a una posibilidad, mientras que la razón se usaría en el cálculo de la probabilidad, aparece en la estrategia SI.4.5.
- *Pensar que se pueden comparar muestras de distinto tamaño usando frecuencias absolutas.* Este conflicto también es descrito en Estepa (1993) y Arteaga (2008), quienes indican que se debe a una incomprensión del concepto de distribución, pues los estudiantes tratan de comparar distribuciones usando sólo frecuencias absolutas. También se describe en las investigaciones de Shaklee y Tucker (1980) y Allan y Jenkins (1983) y en nuestro estudio se presenta en las estrategias SP.4.1, SP.4.2, SI.4.4, SI.4.5 y SI.4.6.
- *Confundir frecuencia con valor de la variable.* Esta confusión indica incomprensión del concepto de distribución y fue descrita por Carvalho (2001) en su estudio sobre la comprensión de la media y por Ruiz (2006) en un trabajo con la variable aleatoria. En nuestro trabajo lo encontramos en la estrategia SI.4.3.
- *Calcular la media de las frecuencias.* El estudiante supone que se puede calcular la media de las frecuencias, no comprendiendo el significado de la media, que no tiene sentido en variables cualitativas, error descrito por Estepa (1993) y otras investigaciones posteriores sobre la comprensión de las medidas de posición central como Cobo (2003) y Mayén (2009). Aparece en las estrategias SP.4.3, SP.5.4 y SI.5.8.
- *Confundir variable cuantitativa y cualitativa, es decir la escala de medida de la variable.* Implica incomprensión de la diferencia entre estos tipos de escalas y variables y lo hemos hallado en la estrategia SP.5.4.

- *Confundir frecuencias dobles y condicionales.* Sería similar al descrito por Falk (1986) en relación a la probabilidad condicional, donde los estudiantes confunden probabilidad condicional y conjunta. Aparece en la estrategia SI.5.6 y fue también descrito en el trabajo de Contreras (2011).
- *Confundir $P(B/A)$ con $P(A/B)$,* es decir, intercambiar los términos de una probabilidad condicional, error descrito por Falk (1996) y denominado por el autor *falacia de la condicional transpuesta.* Aparece en los trabajos de Diaz (2007) y Contreras (2011) y en la estrategia SI.5.7.
- *Confundir una tabla de contingencia con una tabla de frecuencias.* Los estudiantes no comprenden la finalidad de la tabla de contingencia, donde se presentan datos de una distribución bivariante y la tratan como si fuese la distribución de una de una sola variable. Por tanto están confundiendo variable doble con variable simple y frecuencia doble con frecuencia marginal. Aparece en la estrategia SI.5.8. No lo hemos encontrado descrita en los trabajos previos.
- *Obtener una probabilidad mayor que 1.* El estudiante llega a este resultado al confundir casos favorables con posibles y además no es consciente de un axioma de la probabilidad (la probabilidad es un número entre 0 y 1). Este error es descrito en Contreras (2011). Aparece en la estrategia SI.3.5.

Otros conflictos:

- *Basar el juicio de asociación en las teorías previas,* en lugar de utilizar los datos. Aparece en la estrategia SI.0.1. Son los estudiantes que manifiestan la correlación ilusoria descrita por Chapman y Chapman (1969), usando sus teorías sobre la relación entre variables y no evaluando correctamente las contingencias empíricas. Aparece en la estrategia SI.0.1.
- *No usar los datos.* Estudiantes que responden al problema sin utilizar los datos del enunciado. Aparece en las estrategias de nivel 0.
- *Cálculos incorrectos.* Tiene menor importancia, pero, en todo caso, supone un conflicto de aplicación o desarrollo de procedimientos presente en muchas investigaciones (Estepa, 1993; Contreras, 2011). Aparece en la estrategia SI.1.1.

Observamos que se repiten los descritos por Estepa (1993), especialmente los que denomina concepción local, causal y unidireccional de la asociación. Nosotros no

hemos podido ver la estabilidad de estas concepciones, por eso usamos el término conflicto que nos parece menos fuerte que el de concepción, pues en principio sólo implica una confusión por parte del estudiante, más que una idea asumida con convicción. También aparecen nuevos conflictos no descritos por este autor.

3.11. CONCLUSIONES SOBRE EL ESTUDIO INICIAL DE EVALUACIÓN

El objetivo principal del estudio descrito en este capítulo fue realizar un estudio de evaluación sobre las estrategias intuitivas, juicios de asociación y estimación de la intensidad de la asociación en tablas contingencia en los estudiantes de psicología.

Para ello se comenzó seleccionando y adaptando algunos ítems del cuestionario usado en la investigación de Estepa (1993), en concreto, cuatro de los relacionados con las tablas de contingencia. Se realizaron algunos cambios en las variables de tarea y, después de probar el instrumento piloto, se ha llevado a cabo un estudio detallado con una muestra de 414 estudiantes de tres universidades españolas. Además de comparar con los resultados previos de Estepa y otros autores, se completa con el análisis semiótico de las estrategias y la identificación y clasificación de conflictos semióticos. Para finalizar el capítulo, en esta sección presentamos las principales conclusiones de este estudio.

Los resultados indican que los estudiantes consideran mayoritariamente la existencia de asociación en todos los ítems, con mayor frecuencia en los casos en que existe relación en los datos, pero también en el ítem de independencia perfecta. Hemos interpretado este hecho por la influencia del fenómeno denominado correlación ilusoria (Chapman y Chapman, 1969). El 67% indica asociación en este ítem, mientras que en el estudio de Estepa (1993) sólo el 31,4% de estudiantes indican que hay asociación en el caso de independencia perfecta, aunque los datos numéricos del ítem son exactamente los mismos en las dos investigaciones.

La estimación del coeficiente de asociación mejora si comparamos con los resultados de Estepa (1993), en todos los ítems salvo en este caso de independencia. Es precisamente en este caso donde mayor influencia tiene el uso de estrategias formalmente correctas para mejorar la estimación del coeficiente. Los resultados son similares en las tres universidades participantes, observándose asimismo la alta consistencia entre juicio emitido y valor estimado del coeficiente, siendo la estimación más intensa cuando el juicio de asociación es positivo que cuando es negativo. La influencia de la corrección de la estrategia sobre la precisión de la estimación no fue

muy clara, salvo en el ítem 1 (independencia), donde se observa una clara influencia. Atribuimos este resultado al hecho de que algunas estrategias parcialmente correctas han funcionado bien en algunos ítems y también por las pocas diferencias en el valor estimado en los diferentes ítems salvo en el primero.

Los estudiantes expresan pocas teorías previas sobre la asociación, salvo en el ítem de independencia, donde se presenta el mayor número de estudiantes que expresan una teoría previa (9,7%), posiblemente por ser la única pregunta que las teorías previas van en contra de los datos. Los resultados son compatibles con los de Estepa (1993).

El estudio también indica unos pobres resultados en las estrategias para realizar juicios de asociación en los estudiantes que ingresan en Psicología. Al comparar con el estudio de Estepa (1993), realizado con estudiantes del Curso de Orientación Universitaria se muestran en nuestro caso un menor porcentaje de estrategias parcialmente correctas para los cuatro ítems del cuestionario y las tres estrategias más frecuentes en ambos estudios fueron incorrectas, siendo en los dos casos la más frecuente “utilizar una única distribución condicional” que indica una concepción local de la asociación.

Menos del 20% de los estudiantes alcanzan el nivel 5, en la clasificación de Pérez Echeverría (1990), es decir, utilizan todos los datos y los comparan con métodos multiplicativos. La frecuencia de dicho nivel es menor que la encontrada por Estepa, excepto en el último ítem; los niveles más frecuentemente utilizados fueron usar una distribución condicional (nivel 2), y estrategias aditivas con todos los datos (nivel 4), aproximadamente por la tercera parte de estudiante en cada uno de estos niveles y ligeras variaciones en los ítems.

En nuestro trabajo encontramos algunas nuevas estrategias no descritas en la investigación previa. Por otro lado, el análisis de los conflictos semióticos realizado en este estudio confirma la presencia de las concepciones local, causal y unidireccional de la asociación descritas por Estepa en una proporción importante de estudiantes, quienes también en ocasiones se guían por sus teorías previas, mostrando la correlación ilusoria. Además se ofrece una clasificación de estos conflictos, dependiendo de si se relacionan con la confusión de conceptos o respecto a la atribución de propiedades incorrectas a la asociación e independencia. Dicha clasificación amplía la variedad de dificultades descritas por Estepa y al mismo tiempo da una explicación menos exigente de las mismas, pues el conflicto semiótico supone un error de interpretación y no tiene el carácter de estabilidad y resistencia al cambio descrito en las concepciones.

Los resultados de este primer estudio indican la necesidad de continuar la investigación didáctica al respecto. Puesto que la concepción causal y la correlación ilusoria no parecen mejorar con una enseñanza tradicional (Batanero, Estepa y Godino, 1997) para continuar el trabajo se diseñó una propuesta alternativa de enseñanza, donde hemos tratado de incluir actividades que tengan en cuenta los conflictos señalados, así como las variables de las tareas que influyen en los juicios de asociación.

Esta propuesta se describe en el Capítulo 4 y fue experimentada en el curso 2010-2011 con estudiantes de psicología. Los resultados de la evaluación del aprendizaje se describen en los capítulos 5 y 6.

CAPÍTULO 4.

DISEÑO DE UN PROCESO DE ESTUDIO DE LAS TABLAS DE CONTINGENCIA PARA ESTUDIANTES DE PSICOLOGÍA

- 4.1. Introducción
- 4.2. Descripción general del proceso de estudio
 - 4.2.1. Contexto curricular
 - 4.2.2. Estructuración
 - 4.2.3. Teoría sobre la instrucción implícita en el diseño
 - 4.2.4. Distribución temporal
- 4.3. Análisis de la primera lección. Tablas de contingencia
 - 4.3.1. Objetivos y contenidos
 - 4.3.2. Desarrollo
 - 4.3.3. Significado institucional pretendido
- 4.4. Análisis de la segunda lección. Asociación estadística, dependencia funcional e independencia
 - 4.4.1. Objetivos y contenidos
 - 4.4.2. Desarrollo
 - 4.4.3. Significado Institucional pretendido
- 4.5. Análisis de la tercera lección. Estadístico Chi-cuadrado y contrastes asociados
 - 4.5.1. Objetivos y contenidos
 - 4.5.2. Desarrollo
 - 4.5.3. Significado institucional pretendido
- 4.6. Análisis de la cuarta lección. Medidas de asociación
 - 4.6.1. Objetivos y contenidos
 - 4.6.2. Desarrollo
 - 4.6.3. Significado institucional pretendido
- 4.7. Sesión de prácticas
 - 4.7.1. Programa de cálculo Excel
 - 4.7.2. Ejercicios de la primera sesión de prácticas
 - 4.7.3. Significado institucional pretendido
- 4.8. Conclusiones sobre el significado institucional pretendido

4.1. INTRODUCCIÓN

En este capítulo se analiza el diseño de un proceso de estudio sobre las tablas de contingencia, dirigido a la formación de estudiantes de Psicología. En la determinación del significado institucional de referencia (Capítulo 1), que sería la base del diseño, se tuvo en cuenta el análisis del tema en diversos libros de estadística orientados a Psicología o Educación (por ejemplo, Ruiz-Maya, Martín, Montero y Uriz, 1995; Ato y López, 1996; Lipschutz y Schiller, 1999; Ruiz-Maya y Martín, 2002; Batanero y Díaz, 2008).

A partir de dicho análisis, se seleccionaron los objetos matemáticos más adecuados a la construcción de la propuesta didáctica, organizándolos en un proceso de estudio pretendido. Se contextualizaron los diferentes objetos utilizando ejemplos relacionados con la psicología, tomados de los textos de Siegel (1983), Cuadras, Echeverría, Mateo. y Sánchez (1984), San Martín y Pardo (1989), Amón (1993), Martín y Luna (1994), Sánchez (1995), Pardo y San Martín (1998), Lubin, Maciá y Rubio (2005), Solanas, Salafranca, Fauquet y Núñez (2005) y Guàrdia, Freixa, Però y Turbany (2007).

Con ello se pretendía aumentar la motivación del estudiante, al ver la aplicabilidad del tema en su profesión, pues la motivación influye directamente en el rendimiento (Gundlach, Kuntze, Engel y Martignon, 2010). Todo ello se secuenció en orden de dificultad progresiva, en base al tiempo disponible. También se tuvieron en cuenta las directrices de la ANECA (2010) y del Espacio Europeo de Educación Superior para el Título de Grado en Psicología.

En la línea actual señalada en estas directrices para la docencia universitaria, se incorporó la tecnología a la acción didáctica. Se pensó en el uso de Internet, mediante una página Web (<http://www.ugr.es/~analisisdedatos/webcurso/presentacion.html>), donde se depositó el material didáctico al que los estudiantes podrían acceder cuando lo necesitasen. Se utilizaría en el aula el ordenador y proyector, pues se prepararon presentaciones de PowerPoint, tanto para el desarrollo de la teoría, como para el planteamiento y resolución de los ejercicios. Incluso, se preparó un programa en Excel, que implementa todos los cálculos y gráficos requeridos, con el que trabajarían los estudiantes en la sala de informática. Finalmente se implementó una tutoría a distancia para resolver las dudas de los estudiantes.

En lo que sigue, se analiza la secuencia didáctica planificada, lo que permitirá dar una primera evaluación de su *idoneidad didáctica*, utilizando los criterios propuestos por Godino, Contreras y Font (2006) para el análisis de procesos de estudio de las matemáticas, que se completará en el capítulo de conclusiones.

4.2. DESCRIPCIÓN GENERAL DEL PROCESO DE ESTUDIO

4.2.1. CONTEXTO CURRICULAR

La experiencia de enseñanza se llevaría a cabo en la asignatura de *Técnicas de análisis en la investigación psicológica*, dentro del primer curso del Grado de

Psicología, que cuenta con 6 créditos (ECTS), y tiene carácter obligatorio. Esta asignatura se imparte en el Segundo cuatrimestre de cada curso académico. Un prerrequisito de dicha asignatura es tener superada la asignatura: *Descripción y exploración de datos en Psicología*, con 6 créditos (ECTS), también obligatoria y de primer cuatrimestre; se recomienda también que los estudiantes tengan conocimientos adecuados de cálculo aritmético y algebraico elemental y de informática a nivel de usuario. A continuación describimos con más detalle estas dos asignaturas.

Asignatura: Descripción y exploración de datos en Psicología

En la Memoria de verificación del grado de Psicología (ANECA, 2010), los descriptores de la asignatura son: “*descripción de datos, modelos probabilísticas de procesos psicológicos, muestreo, inferencia y generalización*” (p. 38).

Los objetivos (expresados como resultados esperables de la enseñanza) son: (a) conocer los tipos de variables psicológicas y escalas de medida, (b) conocer la descripción y exploración de un grupo de datos, (c) conocer los principios básicos del muestreo, y (d) conocer algunos conceptos básicos de Inferencia Estadística.

Las competencias generales y específicas previstas son las siguientes: (a) Ser capaz de identificar los tipos de variables psicológicas, (b) ser capaz de diferenciar entre niveles de medida, (c) saber realizar la descripción y exploración de un grupo de datos, (d) identificar los modelos probabilísticos que se utilizan en Psicología, (e) reconocer los principios básicos del muestreo y la inferencia Estadística, y (f) interpretar la significación estadística. Para conseguir estos objetivos y competencias, se incluyen los siguientes contenidos:

- Introducción a la teoría de la medición.
- Los datos como resultado de una medición en Psicología, tipos de datos, descripción y exploración de un grupo de datos.
- Descripción multivariante de un grupo de datos, introducción a los modelos probabilísticos discretos y continuos.
- El muestreo en la planificación de las investigaciones, generalización e inferencia.
- El problema de la estimación de parámetros en el análisis de datos de investigación, formulación y contrastación de hipótesis (significación estadística frente a relevancia).

Asignatura: Técnicas de análisis en la investigación psicológica

Según memoria de verificación del grado (ANECA, 2010), esta asignatura tiene los siguientes objetivos: (a) conocer los modelos estadísticos para la realización de inferencia, (b) conocer y aplicar las técnicas de análisis asociadas a los diferentes tipos de diseños de investigación en Psicología, (c) conocer el análisis de los datos en la evaluación de programas, y (d) aprender a analizar datos en programas e intervenciones psicológicas.

Las competencias generales y específicas a desarrollar son: (a) Realizar inferencias estadísticas, reconocer y aplicar las técnicas de análisis asociadas a los diferentes tipos de diseños de investigación en Psicología; (b) Identificar los modelos de análisis multivariante en investigaciones psicológicas, y ser capaces de interpretar los resultados tras el análisis en programas e intervenciones psicológicas.

Para lograrlos se tienen en cuenta los contenidos que se exponen a continuación:

- Formulación y contrastación de hipótesis, contrastes paramétricos;
- Formulación y contrastación de hipótesis, contrastes no paramétricos;
- Inferencias en modelos de regresión; modelos con variables independientes categóricas; modelos con variables independientes continuas y categóricas; modelos con errores dependientes;
- Análisis multivariante en la investigación psicológica.

En consecuencia, los estudiantes potenciales que participasen en el proceso de estudio pretendido tendrían conocimientos de estadística descriptiva, cálculo de probabilidades, variables aleatorias, distribuciones de probabilidades clásicas, distribuciones bivariadas, y contrastes de hipótesis, además de haber trabajado previamente con Excel.

4.2.2. ESTRUCTURACIÓN

La enseñanza se aplicaría en el segundo cuatrimestre académico en los meses de Abril y Mayo, momento que corresponde con el final de la asignatura, pues, aunque posteriormente se realizan clases de refuerzo, en el mes de Junio no se introduce temario nuevo.

El proceso de estudio se organizaría en cuatro lecciones, que se centran en los siguientes temas y que se analizarán con detalle en el resto del capítulo:

1. Tablas de contingencia, lectura e interpretación.
2. Asociación estadística, dependencia funcional e independencia.
3. El estadístico Chi-cuadrado y contrastes asociados.
4. Medidas de asociación.

Estas lecciones forman parte de los bloques de contenido: inferencias en modelos de regresión; modelos con variables independientes categóricas; modelos con variables independientes continuas y categóricas; de la asignatura “*Técnicas de análisis en la investigación psicológica*”

Las lecciones incluyen objetivos, contenidos teóricos, ejemplos y ejercicios resueltos, todo ello desarrollado en el material didáctico que se incluye en el Anexo 5 (Desarrollo teórico de los temas) y que fue entregado impreso a los estudiantes, además de ser colocado en la página web preparada al efecto. También se prepararon diferentes pruebas de evaluación, que se describe con más detalle en los Capítulos 5 y 6, para ser utilizadas en tres momentos del proceso de estudio:

- Se propondría la resolución de un total de 5 problemas abiertos, en tres sesiones, que se alternarían con las exposiciones teóricas, una en el aula de informática (tres problemas) y dos en el aula tradicional (1 problema cada día junto con el cuestionario que se comenta en el siguiente punto). En la primera sesión en el aula de informática se explicaría el funcionamiento del programa, mientras que en la segunda se les pediría que trabajasen sin ayuda del profesor ó de los apuntes.
- Se propusieron dos cuestionarios con preguntas con varios subítems de verdadero/falso. Correspondiendo el primer cuestionario a las lecciones 1 y 2, y el segundo cuestionario a las lecciones 3 y 4.
- En la evaluación final de la asignatura se incluirían 4 preguntas de opción múltiple sobre los contenidos del proceso de estudio.

4.2.3. TEORÍA SOBRE LA INSTRUCCIÓN IMPLÍCITA EN EL DISEÑO

En el proceso de estudio diseñado se tuvo en cuanto la dimensión instruccional de nuestro marco teórico (Godino, 1999; Godino, Contreras y Font, 2006), y en particular, el concepto de trayectoria didáctica, que permite dar cuenta de las interacciones entre el conocimiento matemático, las funciones docentes y discentes, y los procesos semióticos que se ponen en juego en la instrucción matemática.

Los autores definen la trayectoria didáctica como la conjunción interactiva entre las trayectorias epistémica, docente y discente relativas a un contenido y a unas circunstancias dadas. Una trayectoria docente (o discente), es la secuencia de funciones docentes (o discentes) que se generan a lo largo del proceso de estudio de un contenido. Dichas funciones, la mayoría de las cuáles también aparecen en el trabajo de Tauber (2001) de describen a continuación.

- *Funciones docentes:* La intención fue que el profesor controlase el proceso de estudio y los significados presentados en cada sesión. A lo largo del proceso de estudio *el profesor* haría cuatro presentaciones, donde resumiría el material escrito previamente entregado a los estudiantes. Se esperaba que motivase los diferentes lenguajes, conceptos, propiedades, procedimientos y argumentos, mostrando su utilidad en el análisis de los problemas que dan origen a las tablas de contingencia. Presentaría algunos ejemplos y ejercicios resueltos, la mayoría de ellos en contextos de aplicación de psicología y el resto en contextos familiares para los estudiantes. Enseñaría el manejo del programa de cálculo y atendería a los estudiantes durante las sesiones prácticas. También se esperaba que resolviese sus dudas, desde el final del proceso de estudio, hasta el examen final de la asignatura a través de una tutoría virtual por correo electrónico. Finalmente llevaría a cabo la evaluación con una finalidad investigadora, orientada a introducir posibles cambios para la mejora del proceso de estudio.
- *Funciones discentes:* *Los estudiantes*, deberían asistir a todas las sesiones y estudiar los apuntes entregados. Se esperaba también su participación en los distintos tipos de situaciones didácticas en el sentido de Brousseau (1997): (a) Ante los problemas planteados deberían realizar una actividad de exploración y búsqueda de conjeturas y soluciones (situaciones de acción); (b) Formularían y comunicarían sus soluciones (situación de formulación/ comunicación); (c) Participarían en debates sobre las posibles soluciones (situaciones de validación) y (d) Aceptarían el conocimiento resumido por el profesor (situaciones de institucionalización), planteando las dudas sobre los aspectos que no comprendieran, mediante preguntas en clase o preguntas por correo electrónico. Se esperaba que los estudiantes intentasen resolver, individualmente las actividades planteadas y las actividades de evaluación.

En el diseño de la enseñanza también se tuvieron en cuenta los principios de diseño instruccional propuestos por Cobb y McClain (2004) para el aprendizaje de la

estadística que Garfield y Ben-Zvi (2008) describen como “entorno de aprendizaje del razonamiento estadístico”:

1. Énfasis en el desarrollo de las ideas estadísticas fundamentales, en lugar de presentar la estadística como un conjunto de procedimientos. Teniendo en cuenta este principio se trató de relacionar los principales objetos estadísticos en forma progresiva y sin excesiva formalidad;
2. Uso de situaciones reales y motivadoras que sirvan a los estudiantes para realizar y evaluar conjeturas y apoyen el desarrollo de su razonamiento. Para ello se contextualizaron los ejemplos y ejercicios, bien en el campo de la Psicología, o en situaciones familiares para el estudiante;
3. Integrar la tecnología, en forma que permita a los estudiantes evaluar sus conjeturas, explorar y analizar datos y desarrollar su razonamiento estadístico. Por este motivo se proporcionó a los estudiantes un programa de cálculo, liberando así tiempo para actividades de interpretación y exploración;
4. Promover el debate en el aula, favoreciendo el intercambio de las ideas y argumentos de los estudiantes y el descubrimiento dirigido;
5. Uso de la evaluación para informar sobre lo que los estudiantes aprenden, apoyarles en el aprendizaje y comprobar que se alcanzan los objetivos educativos.

4.2.4. DISTRIBUCIÓN TEMPORAL

Se planificaron 6 sesiones de 1 hora de duración como tiempo dedicado a la enseñanza. Cuatro de dichas sesiones se llevarían a cabo en el aula tradicional en grupo mediano (entre 45 y 55 estudiantes por grupo) y se dedicarían a la presentación de los temas, con ayuda de las diapositivas de PowerPoint preparadas al efecto. Al estudiante se le entregaría el material didáctico preparado, que se incluye en el Anexo 5.

También se planificaron dos sesiones prácticas en el laboratorio de informática en las cuáles cada estudiante trabajaría independientemente con el ordenador, con apoyo de un programa Microsoft Excel, preparado para realizar las actividades prácticas (véanse Anexos 6 y 7). En las sesiones prácticas cada grupo de teoría se dividiría en tres subgrupos de aproximadamente 15 estudiantes. En la Tabla 4.2.1 se recoge el calendario de las sesiones que se desarrollarían a lo largo de tres semanas. Puesto que el número total de estudiantes era alrededor de 100, habría dos grupos para las sesiones de teoría y seis para las de prácticas; todos ellos impartidos por el autor de la tesis, que actuaría como profesor.

Para asegurar la validez de la recogida de datos, las clases fueron observadas. Dos observadores miembros del equipo de investigación colaboraron en la observación; cada uno de ellos observó un grupo completo de los dos disponibles. 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. Un resumen de la observación se incluye en el Anexo 4. Los profesores habituales de los cursos también asistieron a las sesiones.

Tabla 4.2.1. Calendario de las sesiones

	Martes Semana 1	Miércoles Semana 1	Martes Semana 2	Miércoles Semana 2	Martes Semana 3	Miércoles Semana 3	Martes Semana 4	Miércoles Semana 4
Grupo 1	Tema 1	Tema 2	Tema 3	Tema 4			Evaluación Temas 1 y 2	Evaluación Temas 3 y 4
			Prácticas 1 y 2		Prácticas 3 y 4			
Grupo 2	Tema 1	Tema 2	Tema 3	Tema 4			Evaluación Temas 1 y 2	Evaluación Temas 3 y 4
				Prácticas 1 y 2		Prácticas 3 y 4		

A continuación se analiza cada una de las lecciones previstas cuyos objetivos emanan de los objetivos de la asignatura “*Técnicas de análisis en la investigación psicológica*”, descritos en la sección 4.2.1.

Las competencias que se trata de lograr son las ya descritas para la citada asignatura, concretadas al estudio de las tablas de contingencia, esto es, (a) Realizar inferencias estadísticas sobre tablas de contingencia, reconocer y aplicar las técnicas de análisis asociadas a dichas tablas en la investigación en Psicología; (b) Identificar los modelos de análisis de tablas de contingencia en investigaciones psicológicas, y ser capaces de interpretar los resultados tras el análisis en programas e intervenciones psicológicas.

Puesto que estos objetivos y competencias son los mismos en las cuatro lecciones, no se repiten en lo que sigue.

4.3. ANÁLISIS DE LA PRIMERA LECCIÓN. TABLAS DE CONTINGENCIA: LECTURA E INTERPRETACIÓN

4.3.1. OBJETIVOS Y CONTENIDOS

En la primera lección se pretende que los estudiantes aprendan a:

- Resumir datos sobre dos variables estadísticas en una tabla de contingencia;
- Identificar las frecuencias que corresponde a cada par de valores de las variables;
- Calcular las frecuencias relativas dobles e interpretarlas;
- Calcular las frecuencias marginales, e interpretarlas;
- Calcular las frecuencias condicionales e interpretarlas;
- Representar gráficamente los datos mediante diagrama de barras adosadas, diagrama de barras apiladas y gráfico tridimensional; y
- Calcular probabilidades simples, compuestas y condicionales a partir de datos de una tabla de contingencia.

En este tema se introducen las tablas de contingencia y los diferentes tipos de frecuencias, así como sus relaciones. Se definen las distribuciones condicionales y marginales y las representaciones gráficas más frecuentes.

Sus contenidos se describen con detalle en el análisis del desarrollo (Sección 4.3.2) y se presentan organizados respecto al tipo de objeto matemático considerado en nuestro marco teórico en la Tabla 3.4.8. Se organiza en los siguientes apartados, todos ellos precedidos y seguidos por ejercicios y ejemplos que permiten contextualizar el tema.

4.3.2. DESARROLLO

Se comenzaría explicando a los estudiantes que, hasta este momento, habrían estudiado las variables unidimensionales, que en estadística no sólo interesa el estudio de variables aisladas, sino muchas veces interesa analizar la variación conjunta de dos variables relativas a la misma unidad estadística, que constituyen las variables bidimensionales. Se seguiría planteando el Ejemplo 1.1 sobre problemas neuronales y efecto de un fármaco, que, a nuestro entender, motivaría la introducción del tema.

Ejemplo 1.1. Problemas neuronales

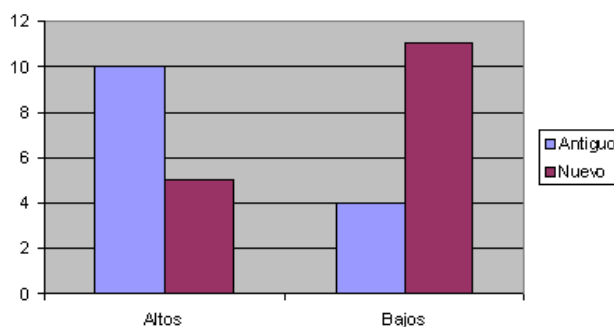
En un hospital psiquiátrico se hace un estudio en el que participan 30 pacientes con dos tipos de problemas neuronales (altos y bajos), queremos comparar un fármaco nuevo con otro antiguo. ¿Cómo podemos representar esta situación? ¿Cómo podemos ver si el tratamiento nuevo es preferible al anterior?

Se indicaría a los estudiantes que, para representar la situación se usarían dos variables estadísticas: la *variable X*, que representaría el tratamiento, con dos valores (x_1, x_2) y la *variable Y* para hacer referencia a los problemas neuronales (y_1, y_2). Se indicaría que el tipo de tabla mostrado en el ejemplo (Tabla 4.3.1) corresponde a la *tabla de contingencia*, más concretamente, de una tabla 2x2. A continuación se representarían los datos en un diagrama de barras adosado (que para el ejemplo sería la Figura 4.3.1).

Tabla 4.3.1. Tipo de tratamiento en enfermos con problemas neuronales

Tipo de tratamiento(X)	Problemas neuronales (Y)	
	Altos(y_1)	Bajos(y_2)
Antiguo(x_1)	10	4
Nuevo(x_2)	5	11

Figura 4.3.1. Diagrama de barras adosadas



Siguiendo a Batanero y Díaz (2008), se introducirían la notación y concepto de frecuencia absoluta doble, destacando que cada uno de los diferentes pares de valores (x_i, y_j) pertenece a la misma unidad estadística (el mismo enfermo), realizando preguntas a algunos estudiantes para que interpretasen las frecuencias de la Tabla 4.3.1. Se introduciría la notación para las frecuencias relativas absolutas dobles f_{ij} y relativas dobles h_{ij} y la fórmula $h_{ij} = \frac{f_{ij}}{n}$ que relaciona los dos tipos de frecuencias. Se calcularían e interpretarían las frecuencias relativas dobles para el ejemplo, resultando la Tabla 4.3.2.

Tabla 4.3.2. Frecuencias relativas dobles

		Problemas neuronales (Y)	
		Altos(y_1)	Bajos(y_2)
Tipo de tratamiento(X)	Antiguo(x_1)	0,333(h_{11})	0,133(h_{12})
	Nuevo(x_2)	0,167(h_{21})	0,367(h_{22})

Se indicaría que, a partir de la tabla de contingencia, pueden obtenerse diferentes distribuciones de una variable: las frecuencias absolutas y relativas marginales por filas y columnas, calculándose e interpretándose para el ejemplo. Se haría observar que la distribución marginal es suma de las frecuencias dobles, en el ejemplo $f_{11}+f_{12}=f_{1.}$ y $f_{21}+f_{22}=f_{2.}$. También se observaría que, al sumar la columna del total por filas, o sumando el total por columnas se obtiene el total de la muestra: $f_{1.}+f_{2.} = f_{.1}+f_{.2} = n$.

A continuación se introducirían las distribuciones condicionales, motivando el tema, e indicando que, en el ejemplo, sería posible centrarse en solamente una parte de los pacientes, por ejemplo en los pacientes con el tratamiento nuevo. Se introduciría para ello las *distribuciones condicionales*. Se introduciría la notación $h(x_i|y_j)$ para la frecuencia relativa condicional del valor x_i entre los individuos que presentan el carácter y_j , y su cálculo (Amón, 1993):

$$h(x_i | y_j) = \frac{f_{ij}}{f_{.j}} = \frac{h_{ij}}{h_{.j}}$$

Igualmente se propondría la notación y cálculo de la frecuencia relativa condicional:

$$h(y_j | x_i) = \frac{f_{ij}}{f_{i.}} = \frac{h_{ij}}{h_{i.}}$$

Se obtendrían las distribuciones condicionales por filas y columnas en el ejemplo obteniendo las Tablas 4.3.3 y 4.3.4, a partir de las cuáles se harían preguntas como, por ejemplo, si hay el mismo porcentaje con problemas neuronales altos y bajos en el nuevo tratamiento.

Tabla 4.3.3. Distribución de problemas neuronales en enfermos con tratamiento nuevo

Tipo de tratamiento(X)	Problemas neuronales (Y)		TOTAL
	Altos(y_1)	Bajos(y_2)	
Nuevo(x_2)	0,3125 ($h(y_1 \text{nuevo})$)	0,6875 ($h(y_2 \text{nuevo})$)	1

Tabla 4.3.4. Distribución de X condicionada a Y₂

Tipo de tratamiento(X)	Problemas neuronales (Y)	
		Bajos(y ₂)
Antiguo(x ₁)		0,2667 ($h(x_1/Y=y_2)$)
Nuevo(x ₂)		0,7333 ($h(x_2/Y=y_2)$)
Total		1

Tabla 4.3.5. Tabla de contingencia para el caso general

Variable X		Variable Y			TOTAL
		Y ₁	...	Y _c	
X ₁		f ₁₁	...	f _{1c}	f _{1.}
...	
X _r		f _{r1}	...	f _{rc}	f _{r.}
TOTAL		f _{.1}	...	f _{.c}	N

Se haría ver que la suma de las frecuencias relativas por filas y columnas es igual a la unidad (Nortes, 1993). Se generalizaría la notación al caso de las tablas rxc (*r* filas y *c* columnas) (Tabla 4.3.5).

Se realizarían otras representaciones gráficas de los datos, mediante el *diagrama de barras apilado* (Figura 4.3.2) y gráfico tridimensional (Figura 4.3.3), en los cuales los estudiantes compararían la información que proporciona cada uno de estos gráficos y el diagrama de barras adosado. También se les indicaría que las gráficas se pueden construir con frecuencias absolutas, relativas o porcentajes. A continuación se introduciría el Ejercicio 1.1, que, además de estar resuelto en los apuntes, se iría haciendo paso a paso en la clase.

Figura 4.3.2. Diagrama de barras apilado

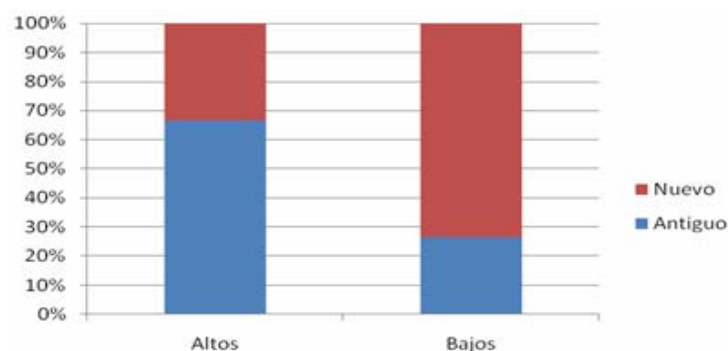
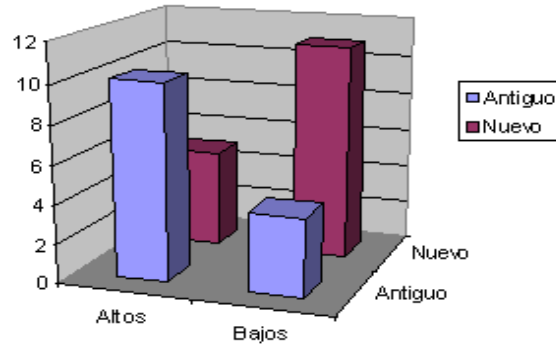


Figura 4.3.3. Gráfico tridimensional



Ejercicio 1.1. Despersonalización

La despersonalización es una alteración de la percepción o la experiencia de uno mismo de tal manera que uno se siente "separado" de los procesos mentales o del cuerpo, como si uno fuese un observador externo a los mismos. La forma severa se encuentra en la ansiedad o en ataques de pánico. En la siguiente tabla clasificamos un grupo de pacientes según el nivel de despersonalización y frecuencia de relaciones personales de la persona que lo presenta:

Relaciones personales (X)	Nivel de despersonalización (Y)	
	Bajo(y_1)	Alto (y_2)
Muchas(x_1)	24	22
Pocas(x_2)	19	40

1. Interpreta los datos de la tabla; calcula e interpreta las frecuencias marginales y relativas dobles y marginales.
2. Calcular e interpretar las frecuencias relativas condicionales por filas y columnas. Si se toma un enfermo al azar ¿Es más probable tener muchas relaciones personales en los pacientes con bajo o alto grado de despersonalización? ¿Es más probable tener alto grado de despersonalización en pacientes con pocas o muchas relaciones interpersonales

Se leería la tabla, haciendo observar, por ejemplo, que hay 24 personas con muchas relaciones personales y bajo nivel de despersonalización, interpretando del mismo modo las otras frecuencias dobles. Respecto a las frecuencias marginales, se calcularían, observando, por ejemplo, que hay 46 personas con muchas relaciones personales (24+22) y 59 personas con pocas relaciones personales (19+40). Se calcularían las frecuencias relativas dobles y marginales, interpretándolas; por ejemplo, el 22,85% de personas tienen muchas relaciones y baja despersonalización $(24/105) \times 100$ y el 43,81% de los participantes tienen muchas relaciones personales $(46/105) \times 100$.

Se continuaría con el cálculo de frecuencias relativas condicionales respecto a cada fila y cada columna. Analizándolas se haría ver que es más probable tener alta despersonalización entre las personas que tiene pocas relaciones personales que entre las que tienen muchas y que los porcentajes en cada fila suman 100. Igualmente se interpretarían los resultados por columnas. Finalmente se mostraría en la Tabla 4.3.6 todos los resultados anteriores.

Tabla 4.3.6. Solución del ejercicio 1.1

Relaciones personales	Despersonalización			Interpretación
	Bajo	Alto	TOTAL	
Muchas	24	22	46	Frecuencia absoluta
	0,23	0,21		Frecuencias relativas
	0,52	0,48		Respecto a fila
	0,56	0,35	0,4381	Respecto a columna
Pocas	19	40	59	Frecuencia absoluta
	0,18	0,38		Frecuencias relativas
	0,32	0,68		Respecto a fila
	0,44	0,65	0,5619	Respecto a columna
Total	43	62	105	
	0,4095	0,5905		Respecto a fila

Se finalizaría el tema con la resolución del Ejercicio 1.2 sobre idealismo, tomado de Sánchez (1995). El método más eficaz para completar la tabla sería buscar las filas y columnas que únicamente les falte un valor (incluido el valor del total).

Ejercicio 1.2. Idealismo

A continuación mostramos una tabla con el cruce idealismo, medida esta variable a partir del grado de acuerdo-desacuerdo de los entrevistados con la siguiente frase: “La vida sólo tiene sentido cuando una persona se dedica plenamente a una causa o ideal” y filiación política

Filiación política	La vida sólo tiene sentido por un ideal				Total	%
	muy en desacuerdo	más bien desacuerdo	más bien en desacuerdo	muy de acuerdo		
Izquierda			420	73	954	58,1
Centro		106		36	298	
Derecha	20		181		389	23,7
Total	143	598		167	1641	
%	8,7	36,4	46,7	10,2	100	

1. Rellena los huecos de la tabla, con las frecuencias y porcentajes que faltan
 2. Interpreta los resultados obtenidos.
- (Sánchez, 1995, p. 300).

Paso a paso se iría completando la tabla, finalizando con el cálculo del “Total”, llegando a la Tabla 4.3.7. En la segunda parte del ejercicio se realizaría la interpretación de la tabla, indicando por ejemplo, que entre los que están muy de acuerdo el 43% es de izquierdas, mientras que los que están muy en desacuerdo el 69% son de izquierdas.

Tabla 4.3.7. Solución del ejercicio 1.2

Ideología Recodificada	Vida sólo tiene sentido				Total	%
	Muy en Desacuerdo	Más bien Desacuerdo	Más bien desacuerdo	Muy de acuerdo		
Izquierda	99	362	420	73	954	58,1
Centro	24	106	132	36	298	18,2
Derecha	20	130	181	58	389	23,7
Total	143	598	733	167	1641	
%	8,7	36,4	46,7	10,2	100	

4.3.3. SIGNIFICADO INSTITUCIONAL PRETENDIDO

Como resumen del análisis del contenido que acabamos de realizar, se presenta una síntesis del significado institucional pretendido en la lección 1 (Tabla 4.3.8). Observamos que en esta lección se presentan dos campos de problemas: (a) resumir la información sobre datos bivariantes y (b) calcular los diferentes tipos de probabilidades, a partir de los datos de la tabla ya formada. Nos hemos interesado por este segundo campo de problemas, no específico de la asociación, puesto que ha sido difícil en varias investigaciones, por ejemplo en Contreras (2011) con estudiantes de magisterio y también en Díaz (2007) con estudiantes de psicología.

Tabla 4.3.8. Síntesis del significado pretendido en la primera lección

Tipo de objetos	Objetos matemáticos
Situaciones problemáticas	P1. Resumir la información obtenida de un conjunto de observaciones bivariantes de modo que se visualicen posibles relaciones entre las variables.
	P2. Cálculo de probabilidades asociadas a valores específicos de una o las dos variables.
Lenguaje	R2. Diagrama de barras apilado.
	R3. Gráfico tridimensional.
	R4. Lenguaje verbal.
	R5. Lenguaje simbólico.
	R6. Lenguaje gráfico.
	R7. Tablas (de frecuencia y de contingencia).
Procedimientos	PR1. Lectura e interpretación de la tabla.
	PR2. Construcción de una tabla de contingencia.
	PR3. Cálculo de probabilidades simples, compuestas y condicionales a partir de la tabla.
	PR4. Cálculo de frecuencias condicionales y marginales (absolutas o relativas).
	PR5. Representación gráfica de datos a partir de una tabla de contingencia.
Conceptos	C1. Variable simple / doble valor.
	C2. Distribución bidimensional.
	C3. Variable cualitativa y cuantitativa.
	C4. Tablas de contingencia, tablas rxc.
	C5. Frecuencia absoluta y relativa doble.
	C6. Frecuencia absoluta y relativa marginal; distribución marginal.
	C7. Frecuencia relativa condicional; distribución condicional.
	C8. Probabilidad simple.
	C9. Probabilidad conjunta y condicional.
	C14. Diagrama de barras apilado.
	C15. Diagrama de barras adosado.
	C16. Gráfico tridimensional.
Propiedades	PP1. Relaciones entre frecuencias absolutas y relativas.
	PP2. Relaciones entre frecuencias dobles, condicionales y marginales.
	PP6. Suma de frecuencias relativas igual a la unidad.
	PP14. El diagrama de barras adosado se puede construir con frecuencias absolutas o porcentajes.
	PP15. Relación y diferencia entre diagrama de barras adosado y apilado.
	PP21. Relación entre probabilidad y tabla de contingencia.
	PP22. Suma de frecuencias absolutas dobles igual al total de la muestra.
Argumentos	A1. Razonamientos informales.
	A2. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar; comparación de frecuencias.

Se introducirían los diversos lenguajes, gráficos, verbales y simbólicos asociados a la tablas de contingencia y los conceptos, procedimientos y propiedades asociados a la variable doble, su distribución y valores, los diversos tipos de frecuencias y probabilidades calculables de la tabla y las representaciones gráficas. Los argumentos usados por el profesor para presentar el tema, así como los esperados por parte de los estudiantes serían informales.

Además de los ejemplos expuestos se realizaron otras actividades en las clases prácticas que se describen en la Sección 4.7. La evaluación de esta lección se realizó mediante distintos instrumentos. Los resultados de la evaluación y los instrumentos se analizan en los capítulos 5 y 6.

4.4. ANÁLISIS DE LA SEGUNDA LECCIÓN. ASOCIACIÓN ESTADÍSTICA, DEPENDENCIA FUNCIONAL E INDEPENDENCIA

4.4.1. OBJETIVOS Y CONTENIDOS

En la segunda lección se plantearían los siguientes objetivos:

- Diferenciar la asociación estadística, dependencia funcional e independencia.
- Reconocer el tipo de relación entre dos variables de una tabla comparando las frecuencias condicionales.
- Calcular las frecuencias esperadas en caso de independencia.
- Analizar posibles explicaciones de una asociación estadística: relación causal, interdependencia, tercera variable explicativa o asociación espuria.

Contenidos

La lección se dedica a introducir el concepto de asociación estadística, dependencia funcional e independencia, así como al estudio de las propiedades de la independencia y las posibles explicaciones de la asociación. Sus contenidos se describen con detalle en la siguiente sección y se presentan organizados respecto al tipo de objeto matemático considerados en nuestro marco teórico en la Tabla 4.4.6.

4.4.2. DESARROLLO

Se comenzaría recordando a los estudiantes que en el tema anterior estudiaron las tablas de contingencia, sus representaciones gráficas y la interpretación de los distintos tipos de frecuencias. Se motivaría el nuevo tema indicando que, muchos

investigadores están interesados en saber si las dos variables presentadas en una tabla de contingencia están interrelacionadas entre sí. Se indicaría que en el tema se estudiaría la asociación estadística, diferenciándola de la dependencia funcional y de la independencia.

Se recordaría el concepto de dependencia funcional, poniendo el ejemplo de la velocidad, espacio y tiempo. Se haría observar, a partir del ejemplo, que en la dependencia funcional, los valores que toma la variable *Y* (dependiente) quedan determinados, de un modo preciso, por uno y solo un valor de la otra variable, que se considera como independiente (Nortes, 1993).

El concepto de dependencia aleatoria se introduciría a partir del ejemplo de cociente intelectual frente a una nota obtenida en un examen. Se clarificaría con más precisión comentando la tendencia que tendría una nube de puntos en este caso.

La lección continuaría presentando el concepto de asociación, y los diferentes tipos posibles (independencia, asociación parcial y asociación perfecta). Utilizaríamos el Ejemplo 2.1 de trastornos digestivos, adaptado de Estepa (1993) para hablar de asociación, como concepto que extiende la idea de dependencia funcional, e indicar que nos centraríamos en la asociación entre variables cualitativas, y la Figura 4.4.1 para hablar de los distintos tipos de asociación.

Ejemplo 2.1. Trastornos digestivos

Se quiere estudiar si un cierto medicamento produce trastornos digestivos en los ancianos. Para ello se han observado durante un periodo de tiempo a 25 ancianos obteniendo los siguientes resultados.

Sintomatología digestiva según se toma o no una medicina

	Molestias digestivas	No tiene molestias	Total
Toma la medicina	9	8	17
No la toma	7	1	8
Total	16	9	25

Utilizando los datos de la tabla, razona si en estos ancianos, el padecer trastornos digestivos está relacionado con haber tomado o no el medicamento, indica cómo has usado los datos.
(Adaptado de Estepa, 1993, p. 397)

Figura 4.4.1. Diferentes tipos de asociación

a) Independencia total					b) Asociación parcial					c) Asociación perfecta				
	Y ₁	Y ₂	Y ₃	Total		Y ₁	Y ₂	Y ₃	Total		Y ₁	Y ₂	Y ₃	Total
X ₁	10	20	70	100	X ₁	10	80	10	100	X ₁	100			100
X ₂	20	40	140	200	X ₂	80	20	100	200	X ₂			200	200
Total	30	60	210	300	Total	90	100	110	300	X ₃		100		100

Una vez comprendida la asociación perfecta, se estudiaría la independencia, analizando el significado estadístico de este término. Para ello se calcularían las distribuciones condicionales por filas o columnas, indicando a los estudiantes que en el caso de que coincidan estas distribuciones, se puede afirmar que las dos variables son independientes (resaltando que el estudio de la independencia puede realizarse indistintamente con ambos tipos de distribuciones). Se mostraría el Ejemplo 2.2 de insomnio frente a estrés, obteniendo la Tabla 4.4.1 de frecuencias relativas condicionales por columnas y haciendo ver que la variable X es independiente de Y , ya que las distribuciones de frecuencias relativas condicionales son iguales entre si e iguales a la distribución marginal de la variable X .

Ejemplo 2.2. Insomnio y estrés

Se pretende estudiar si el sufrir insomnio tiene relación con los trastornos de estrés. En una muestra de 250 personas observadas se obtuvieron los siguientes resultados:

	Padece estrés	No padece estrés	Total
Tener insomnio	90	60	150
No tener insomnio	60	40	100
Total	150	100	250

Tabla 4.4.1. Insomnio y estrés

	Padece estrés	No padece estrés	Total
Tener insomnio	$90/150=0,6$	$60/100=0,6$	$150/250=0,6$
No tener insomnio	$60/150=0,4$	$40/100=0,4$	$100/250=0,4$
Total	150	100	250

Se continuaría presentando otras propiedades de la independencia (Batanero y Díaz, 2008), para explicarlas de forma individual cada una, haciendo ver a los estudiantes que se van cumpliendo en el ejercicio:

- La frecuencia relativa doble coincidirían al producto de las frecuencias relativas marginales de su fila y su columna: $h_{ij} = h_{i.} \cdot h_{.j}$, para todo i, j :

$$\begin{array}{llll}
 h_{11} = h_{1.} \cdot h_{.1} & 0,36 = 0,6 \times 0,6 & h_{12} = h_{1.} \cdot h_{.2} & 0,24 = 0,6 \times 0,4 \\
 h_{21} = h_{2.} \cdot h_{.1} & 0,24 = 0,4 \times 0,6 & h_{22} = h_{2.} \cdot h_{.2} & 0,16 = 0,4 \times 0,4
 \end{array}$$

- La frecuencia relativa marginal de cualquier valor de Y condicionada por un valor de X , coincidiría para todos los valores de X , es decir Y no depende de X : $h(y_j/x_i) = h_{.j}$. Podríamos observarlo en el ejemplo.

- Para introducir la tercera propiedad se comenzaría definiendo las frecuencias esperadas en caso de independencia, y describiendo su cálculo: $f_{ij} = \frac{f_{i.} \times f_{.j}}{n}$. Si las

variables son independientes, la frecuencia absoluta esperada en cada celda de la tabla coincidirían al producto de las frecuencias absolutas marginales de su fila y columna, dividido por el total de datos, por ejemplo,

$$f_{12} = f_{1.} \cdot f_{.2} / n \quad 60 = 150 \times 100 / 250$$

A continuación se trabajaría con tablas 2x2, y se definiría la dependencia directa e inversa, para observar que celdas nos informan de una dependencia directa, y cuales de una dependencia inversa, usando el Ejemplo 2.4 sobre alergia y vida sedentaria, adaptado de Estepa (1993).

Ejemplo 2.4. Alergia

Se quiere saber si sufrir o no de alergia tiene relación con llevar una vida sedentaria (llevar una vida sin realizar ningún tipo de ejercicio físico). Los datos de 300 sujetos se presentan a continuación.

Forma de vida	Sufre alergia	No sufre alergia
Sedentaria	130	30
No sedentaria	20	120

(Adaptado de Estepa, 1993, p. 398)

Observaríamos que 250 de las 300 personas de la muestra, tendrían a la vez los dos caracteres o no tendrían ninguno. Estas dos celdas (presencia-presencia y ausencia-ausencia) informarían que la asociación en la tabla es directa, según Inhelder y Piaget (1995). Por el contrario, según estos autores, en las otras dos celdas, se daría un solo carácter y el otro no. Serían las celdas favorables a una asociación inversa (sólo 50 casos). Deducimos que, en el ejemplo, la asociación es directa.

Una vez que los estudiantes hubiesen comprendido este punto, el profesor pasaría a analizar las posibles explicaciones de la asociación (Barbancho, 1973), de manera que los estudiantes comprendiesen que si se encuentra asociación entre dos variables, no deberían pensar automáticamente que una de las variables sea causa de la otra, pues se deseaba concienciar a los estudiantes de la posible existencia de la concepción causal de la asociación descrita por Estepa (1993). Se mostraría el Ejemplo 2.5, en el que se relacionarían el número de libros en una casa con el cociente intelectual de los niños de esa casa, donde es claro que hay muchas variables genéticas o educativas que pueden afectarlo, por lo que la posible explicación de la relación sería la existencia de una tercera variable que afecte a las dos estudiadas.

A continuación se introduciría el Ejercicio 2.1 adaptado de Cuadras, Echeverría, Mateo y Sánchez (1984), que, además de estar resuelto en los apuntes, se iría desarrollando en la clase, atendiendo a posibles dudas de los estudiantes.

Ejercicio 2.1. Decisión en un colegio

De los 650 padres del colegio A, 400 están en contra de que los niños vayan a colonias y el resto a favor, mientras que de los 260 padres del colegio B, 100 están a favor de que los niños realicen esta actividad y el resto en contra. Dispuestas las observaciones en una tabla de contingencia de 2 x 2 tenemos:

Muestras	Opinión		Total
	Favor	Contra	
Colegio A	250	400	650
Colegio B	100	160	260
Total	350	560	910

Estudia si hay asociación en estas variables.

(Adaptado de Cuadras, Echeverría, Mateo y Sánchez, 1984, p. 644).

Se propondría en cálculo de las frecuencias relativas condicionales por columnas, que coincidirían, deduciendo, por tanto, las variables son independientes (Tabla 4.4.2) e indicando que también se puede realizar por filas. Después se introduciría el Ejercicio 2.2, tomado de Martín y Luna (1994), que también está resuelto en los apuntes, para el cual las frecuencias relativas condicionales por columnas (Tabla 4.4.3) no coincidirían, por lo que la conclusión sería que las variables no son independientes. Observando las diversas celdas se concluiría que hay dependencia directa pues el número de casos en la diagonal principal es mucho mayor que en la segunda diagonal.

Tabla 4.4.2. Decisión en un colegio

	Favor	Contra	Total
Colegio A	$250/350=0,714$	$400/560=0,714$	$650/910=0,714$
Colegio B	$100/350=0,286$	$160/560=0,286$	$260/910=0,286$
Total	350	560	910

Ejercicio 2.2. Caries en los niños

Para saber si la presentación de caries en niños está asociada con la experiencia de caries en al menos uno de los padres, se tomaron 523 niños de entre 12 y 15 años de edad y se les clasificó según su estado de caries dental (Baja, Normal y Alta) y según la experiencia de caries en sus padres (Baja, Normal y Alta), obteniéndose los datos de la tabla siguiente.

Niño	Padres		
	Baja	Normal	Alta
Baja	142	20	48
Normal	46	108	47
Alta	30	15	67

Estudia si hay asociación en estas variables (Martín y Luna, 1994, p. 375).

Tabla 4.4.3. Caries en niños

Niño	Padres		
	Baja	Normal	Alta
Baja	142/218=0,651	20/143=0,14	48/159=0,302
Normal	46/218=0,211	108/143=0,755	47/159=0,296
Alta	30/218=0,138	15/143=0,105	67/159=0,421
Total	218	143	159

Se finalizaría el tema con la resolución del Ejercicio 2.3, también de Martín y Luna (1994), sobre infarto de miocardio.

Ejercicio 2.3. Infarto de miocardio

Para evaluar si el ejercicio físico contribuye de algún modo a la prevención del infarto de miocardio, se tomaron 133 individuos que lo habían sufrido y otros 133 controles (sin infarto). También fueron clasificados como vigorosos si dedica más de 20 minutos/semana a algún ejercicio físico o no (si dedica menos). Los datos se presentan a continuación:

	No sufrido	Sufrido
No-vigorosos	25	75
vigorosos	35	105
Total	60	180

Estudia la asociación en las variables. Si son independientes comprueba las propiedades de independencia (Adaptado de Martín y Luna, 1994, p. 376).

Siguiendo las sugerencias de Estepa (1993); para analizar la asociación podríamos comparar las frecuencias relativas condicionales por columnas, que coincidirían (Tabla 4.4.4); por lo tanto las variables serían independientes.

Tabla 4.4.4. Infarto de miocardio

	No sufrido	Sufrido	Total
No-vigorosos	25/60=0,417	75/180=0,417	100/240=0,417
Vigorosos	35/60=0,583	105/180=0,583	140/240=0,583
Total	60	180	240

Tabla 4.4.5. Frecuencias esperadas (infarto de miocardio)

	No sufrido	Sufrido	Total
No-vigorosos	25	75	100
Vigorosos	35	105	140
Total	60	180	240

Los alumnos calcularían las frecuencias relativas dobles, comprobando que se cumple $h_{i,j} = h_i \cdot h_j$, para todo i, j , por ejemplo, $h_{22} = h_{2 \cdot} \cdot h_{\cdot 2}$, $0,4375 = 0,5833 \times 0,75$ y las frecuencias esperadas en caso de independencia $e_{i,j} = \frac{f_{i \cdot} \times f_{\cdot j}}{n}$. Las frecuencias observadas

serían iguales a las frecuencias esperadas en caso de independencia, por tanto, llegarían a la conclusión de que las variables son independientes (Tabla 4.4.5).

4.4.3. SIGNIFICADO INSTITUCIONAL PRETENDIDO

Como resumen del análisis del contenido de la Lección 2, se presenta en la Tabla 4.4.6 una síntesis del significado institucional pretendido. Observamos que en esta lección se introduce de nuevo el campo de problemas P1, consistente en resumir la información obtenida de un conjunto de observaciones bivariantes de modo que se visualicen posibles relaciones entre las variables.

Tabla 4.4.6. Síntesis del significado presentado en la segunda lección

Tipo de objetos	Objetos matemáticos
Situaciones-problemas	P1. Resumir la información obtenida de un conjunto de observaciones bivariantes de modo que se visualicen posibles relaciones entre las variables.
Lenguaje	R4. Lenguaje verbal. R5. Lenguaje simbólico. R7. Tablas (de frecuencia y de contingencia).
Procedimientos	PR1. Lectura e interpretación de la tabla. PR4. Cálculo de frecuencias condicionales y marginales (absolutas o relativas). PR6. Cálculo de las frecuencias esperadas. PR7. Comprobar las propiedades de independencia.
Conceptos	C1. Variable simple / doble valor. C2. Distribución bidimensional. C4. Tablas de contingencia, tablas rxc. C5. Frecuencia absoluta y relativa doble. C6. Frecuencia absoluta y relativa marginal; distribución marginal. C7. Frecuencia relativa condicional; distribución condicional. C10. Independencia /Asociación, dependencia funcional. C11. Asociación inversa, directa. C13. Causalidad. C17. Frecuencias esperadas.
Propiedades	PP1. Relaciones entre frecuencias absolutas y relativas. PP2. Relaciones entre frecuencias dobles, condicionales y marginales. PP3. Igualdad /diferencia de frecuencias relativas condicionales en caso de independencia/dependencia. PP4. Igualdad/ diferencia de frecuencias relativas condicionales y marginales en caso de independencia/ dependencia. PP5. Celdas que indican signo de asociación en la tabla 2x2. PP6. Suma de frecuencias relativas igual a la unidad. PP7. Frecuencias esperadas en caso de independencia. PP9. Relación entre asociación y causalidad. PP10. Diferencia entre dependencia aleatoria y funcional. PP16. Propiedades de independencia.
Argumentos	A1. Razonamientos informales. A2. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar; comparación de frecuencias.

Se introducirían nuevos lenguajes verbales y simbólicos principalmente que complementarían los vistos en la lección anterior, y los conceptos y propiedades asociados a las nociones de dependencia e independencia en una tabla de contingencia, los diversos tipos de dependencias y del signo de la asociación en tablas 2x2. Además de reiterar procedimientos empleados en la primera lección, se introducirían algunos procedimientos informales de evaluación de la asociación. Los argumentos usados por el profesor para presentar el tema, así como los esperados por parte de los estudiantes serían informales, aunque se introduce la deducción de algunas propiedades.

Igualmente, en la sesión de prácticas, descrita al final del capítulo, se realizarían ejercicios de cálculo de las frecuencias esperadas en caso de independencia y comprobación de las propiedades de la independencia en tablas con y sin asociación en los datos. Como en el tema anterior, la evaluación descrita en los Capítulos 5 y 6 tiene en cuenta los contenidos y objetivos de esta lección.

4.5. ANÁLISIS DE LA TERCERA LECCIÓN. ESTADÍSTICO CHI-CUADRADO Y CONTRASTES ASOCIADOS

4.5.1. OBJETIVOS Y CONTENIDOS

Se plantearían como objetivos que los estudiantes aprendiesen a:

- Dar una medida de la diferencia entre las frecuencias observadas y las frecuencias esperadas en caso de independencia en una tabla de contingencia;
- Calcular e interpretar el estadístico Chi-cuadrado y sus grados de libertad; manejar las tablas de la distribución para encontrar puntos críticos;
- Recordar los pasos y conceptos básicos en un contraste de hipótesis, así como su correcta interpretación;
- Comprender los pasos para llevar a cabo el contraste de independencia y adquirir competencia para llevarlos a cabo;
- Contrastar la hipótesis de que dos variables en una tabla de contingencia son independientes, decidiendo si el valor Chi-cuadrado es estadísticamente significativo y tomando una decisión sobre el rechazo o de la hipótesis nula;
- Interpretar el nivel de significación correctamente;
- Realizar un contraste de Chi-cuadrado de homogeneidad para comprobar si varias subpoblaciones vienen de la misma población;

- Comprender los supuestos de aplicación del contraste Chi-cuadrado.

Contenidos

Estaría dedicada a introducir el contraste de independencia y el contraste de homogeneidad, explicando su utilidad, y la diferencia entre los dos tipos de contrastes. Se recordarían los pasos necesarios en la realización de un contraste de hipótesis, que los alumnos habían estudiado en el curso anterior y la interpretación correcta de sus resultados. Se estudiaría el cálculo del estadístico Chi-cuadrado, y de sus componentes, así como el significado de los grados de libertad. Se completa el tema con el análisis de las condiciones de aplicación. Estos contenidos se describen con detalle en el análisis del desarrollo (Sección 4.5.2) y se presentan organizados respecto al tipo de objeto matemático considerados en nuestro marco teórico en la Tabla 4.5.6

4.5.2. DESARROLLO

Se comenzaría haciendo un resumen de la lección anterior, estableciendo un vínculo entre los temas 2 y 3, para ayudar a los estudiantes a establecer un hilo conductor del aprendizaje y motivar el tema. Seguidamente se explicaría que en este tema se les enseñaría dos contrastes de hipótesis (contraste de homogeneidad y contraste de independencia), los cuales utilizarían el mismo estadístico (el Chi-cuadrado). Recordemos que los estudiantes ya habrían estudiado contrastes de hipótesis en la asignatura anteriormente, por lo que conocerían el procedimiento y los términos que se iban a utilizar, que, no obstante, les sería repasado.

Comenzaríamos recordando el cálculo de las frecuencias esperadas y su interpretación, ya que son necesarias para obtener el estadístico. Presentaríamos a

continuación el cálculo de Chi-cuadrado de la siguiente forma: $\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$

(Nortes, 1993). Con el estadístico ya presentado observaríamos una serie de propiedades (Martín y Luna, 1994):

- Si todas las frecuencias observadas son iguales a la correspondiente frecuencia esperada, entonces $\chi_{\text{exp}}^2 = 0$,
- Si hay independencia entre las dos variables de la tabla, $\chi_{\text{exp}}^2 = 0$ pues en este caso hay coincidencia entre los dos tipos de frecuencias;
- Cuanto mayor sea la diferencia entre las frecuencias observadas y esperadas en la

tabla, el valor de Chi-cuadrado será mayor. Es decir, a mayor intensidad de la asociación entre las variables, Chi-cuadrado será mayor.

- El valor de Chi-cuadrado siempre es positivo o cero, al ser una suma de cuadrados.
- En general, a mayor número de sumandos, se obtendrá un valor mayor.

Seguiríamos calculando los grados de libertad, los cuales se recordaría que se refieren al número de cantidades independientes que se necesitan en su cálculo, menos el número de restricciones que ligan a las observaciones y el estadístico. Esto resultaría ser igual a $k = (m-1) \times (n-1)$, siendo m el número de filas, y n el número de columnas. Se expresaría esta dependencia de la forma:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \rightarrow \chi^2_{k=(n-1)(m-1); \alpha}$$

El profesor mostraría los cálculos e interpretación en el Ejemplo 3.1 de supervivencia en el Titanic, utilizando datos tomados de Internet.

Ejemplo 3.1. Supervivencia en el Titanic

El 10 de abril de 1912, el Titanic zarpaba con 1317 pasajeros a bordo, ante la admiración de una muchedumbre de curiosos que contemplaban atónitos como aquella mole de acero se alejaba majestuosamente del puerto. Cinco días después los medios de comunicación de todo el mundo se hicieron eco de la increíble noticia: el barco más grande jamás construido yacía a casi cuatro mil metros de profundidad. La Tabla 4.5.1 muestra la distribución de pasajeros, según supervivencia y tipo de pasaje.

Tabla 4.5.1. Distribución de pasajeros en el Titanic según supervivencia y tipo de pasaje

	Sobrevive	No sobrevive	Total
Primera clase	194	128	322
Segunda clase	119	161	280
Tercera clase	138	573	711
Total	451	862	1313

Calcularíamos en el ejemplo las frecuencias esperadas en caso de independencia: (Tabla 4.5.2) y los grados de libertad: $(3-1) \times (2-1) = 2 = k$. A continuación llevaríamos a cabo los cálculos del estadístico Chi-cuadrado (Tabla 4.5.3).

Tabla 4.5.2. Frecuencias esperadas

	Sobrevive	No sobrevive	Total
Primera clase	110,6	211,4	322
Segunda clase	96,2	183,8	280
Tercera clase	244,2	466,8	711
Total	451	862	1313

4.5.3. Cálculo de Chi-cuadrado

	Sobrevive	No sobrevive
Primera clase	$\frac{(194 - 110,6)^2}{110,6} = 62,9$	$\frac{(128 - 211,4)^2}{211,4} = 32,9$
Segunda clase	$\frac{(119 - 96,2)^2}{96,2} = 5,4$	$\frac{(181 - 183,8)^2}{183,8} = 2,8$
Tercera clase	$\frac{(138 - 244,2)^2}{244,2} = 46,2$	$\frac{(573 - 466,8)^2}{466,8} = 24,2$

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 62,9 + 32,9 + 5,4 + 2,8 + 46,2 + 24,2 = 174,4$$

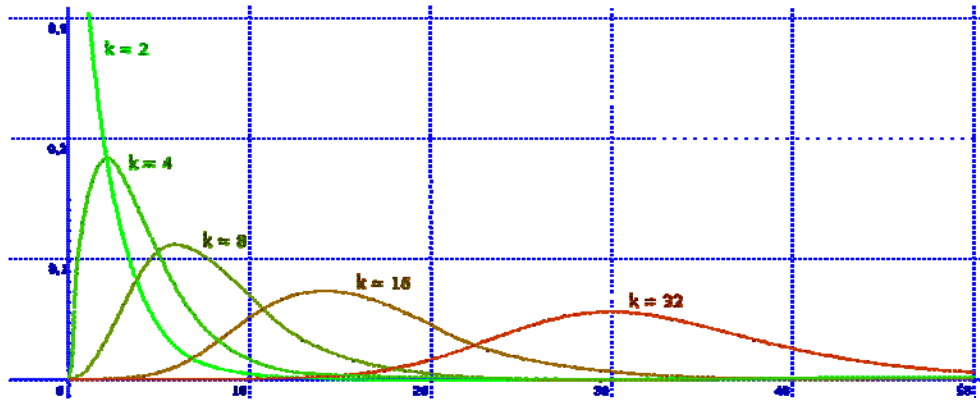
Tabla 4.5.4. Tabla Chi-cuadrado

Grados libertad	Probabilidad de un valor superior - Alfa (α)				
	0,1	0,05	0,025	0,01	0,005
1	2,71	3,84	5,02	6,63	7,88
2	4,61	5,99	7,38	9,21	10,60
3	6,25	7,81	9,35	11,34	12,84
4	7,78	9,49	11,14	13,28	14,86
5	9,24	11,07	12,83	15,09	16,75
6	10,64	12,59	14,45	16,81	18,55
7	12,02	14,07	16,01	18,48	20,28
8	13,36	15,51	17,53	20,09	21,95
9	14,68	16,92	19,02	21,67	23,59
10	15,99	18,31	20,48	23,21	25,19
11	17,28	19,68	21,92	24,73	26,76
12	18,55	21,03	23,34	26,22	28,30
13	19,81	22,36	24,74	27,69	29,82
14	21,06	23,68	26,12	29,14	31,32
15	22,31	25,00	27,49	30,58	32,80
16	23,54	26,30	28,85	32,00	34,27
17	24,77	27,59	30,19	33,41	35,72
18	25,99	28,87	31,53	34,81	37,16
19	27,20	30,14	32,85	36,19	38,58
20	28,41	31,41	34,17	37,57	40,00
21	29,62	32,67	35,48	38,93	41,40
22	30,81	33,92	36,78	40,29	42,80
23	32,01	35,17	38,08	41,64	44,18
24	33,20	36,42	39,36	42,98	45,56
25	34,38	37,65	40,65	44,31	46,93
26	35,56	38,89	41,92	45,64	48,29
27	36,74	40,11	43,19	46,96	49,65
28	37,92	41,34	44,46	48,28	50,99
29	39,09	42,56	45,72	49,59	52,34
30	40,26	43,77	46,98	50,89	53,67

Los alumnos interpretarían el resultado obtenido observando que el valor obtenido 174,4 es muy poco probable en caso de independencia. De ello deducirían que el salvamento de los viajeros en el Titanic no fue independiente de su clase social. El

profesor mostraría la forma que toma el estadístico Chi-cuadrado (Figura 4.5.1) para diferentes valores de su parámetro, recordando a los estudiantes la lectura de la tabla Chi-cuadrado (Tabla 4.5.4) que ya conocen de temas anteriores.

Figura 4.5.1. Distribución Chi-cuadrado



A continuación se introduciría el contraste de independencia, ya empleado en el Ejemplo 3.1, donde los pasos a realizar son los siguientes (Batanero y Díaz, 2008):

1. Fijar las hipótesis (Lipschutz y Schiller, 1999), que es el primer paso en cualquier contraste. En este caso, las hipótesis nula (no efecto) y alternativa serían las siguientes:

H_0 : Las variables en filas y columnas de la tabla son independientes.

H_1 : Hay asociación entre las filas y columnas de la tabla.

2. Fijar el nivel de significación; lo más usual es elegir un valor $\alpha=0,05$, pero hay que recordar a los estudiantes que dicho valor es una elección del investigador y se podría tomar otro diferente (Skipper, Guenter y Nass, 1970):
3. Elegir un estadístico de contraste que nos permita contrastar la hipótesis y tenga relación con ella. En este caso, el estadístico sería el Chi-cuadrado:

$$\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \rightarrow \chi_{(n-1)(m-1)}^2$$

- Se calcularía el valor empírico a partir de los datos. Puesto que, en el caso de independencia esperamos un valor pequeño, si $\chi_{\text{exp}}^2 > \chi_{(n-1)(m-1); \alpha}^2$ este valor sería muy poco probable en caso de independencia. En este caso, tomamos la decisión de rechazar la hipótesis de independencia y por tanto suponemos que las variables están asociadas.

- Si $\chi_{\text{exp}}^2 < \chi_{(n-1) \times (m-1), \alpha}^2$. En este caso no tomamos ninguna decisión, ya que el rechazo de la hipótesis nula tiene más fuerza que su aceptación (Batanero, 2000).

En cualquiera de los dos casos es importante recordar a los estudiantes que se podría cometer un error en la decisión. Mostraríamos lo explicado mediante el ejemplo 3.2 de deporte y bienestar, tomado de Lubin, Maciá y Rubio (2005).

Ejemplo 3.2. Deporte y bienestar

Un investigador quiere estudiar si hay asociación entre la práctica deportiva y la sensación de bienestar. Extrae una muestra aleatoria de 100 sujetos. Los datos aparecen a continuación.

Sensación de Bienestar	Práctica deportiva		Total
	Sí	No	
Sí	20	25	45
No	10	45	55
Total	30	70	100

Contraste la hipótesis de independencia entre bienestar y práctica de deporte (alfa = 0,01). (Lubin, Maciá y Rubio, 2005, p. 324).

Para este ejemplo los estudiantes calcularían el estadístico Chi-cuadrado, que tendría los grados de libertad: $(n-1) \times (m-1) = 1 \times 1 = 1$ y el valor:

$$\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 8,13$$

La lectura de la tabla Chi-cuadrado daría como resultado que la probabilidad de obtener un valor 8,13 o mayor con 1 grado de libertad es $p = 0,004$. Por tanto el valor obtenido del estadístico es estadísticamente significativo, pues la probabilidad de obtenerlo, en caso de ser la hipótesis nula cierta es menor que 0,01. La decisión que se debería tomar es rechazar la hipótesis de independencia entre bienestar y práctica deportiva.

A continuación se introduciría el contraste de homogeneidad, indicando las situaciones donde se utiliza, que son aquellos casos en que se desea comprobar que varias distribuciones de una misma variable en diferentes subpoblaciones son iguales o si hay alguna de ellas que difiere. Se indicaría que el estadístico de contraste y los pasos seguidos son exactamente iguales que los utilizados para el caso del contraste de independencia, variando las hipótesis, que serían las siguientes (Batanero y Díaz, 2008):

- $H_0: p_{1j} = p_{2j} = \dots = p_{mj}$ para todo j ; dicho de otro modo, todas las subpoblaciones tienen idéntica distribución para la variable Y .
- H_1 : algunas de estas proporciones son diferentes. Es decir, la distribución de la variable Y en alguna de estas subpoblaciones es diferente.

Se mostraría el procedimiento completo del contraste de homogeneidad mediante el Ejemplo 3.3 sobre homogeneidad del grupo sanguíneo.

Ejemplo 3.3. Grupo sanguíneo.

Se desea saber si la distribución de los grupos sanguíneos es similar en los individuos de dos poblaciones. Para ello se elige una muestra aleatoria de cada una de ellas, obteniéndose los siguientes datos ¿Qué decisión se debe tomar?

	A	B	AB	0	Total
Muestra 1	90	80	110	20	300
Muestra 2	200	180	240	30	650
Total	290	260	350	50	950

Una vez planteadas las hipótesis, y comprendida la diferencia entre esta situación y el contraste de independencia, se calcularía el estadístico Chi-cuadrado con $(n-1) \times (m-1) = 1 \times 3 = 3$ grados de libertad y valor:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 1,76$$

Se pediría a los estudiantes que buscaran la probabilidad asociada en la tabla Chi-cuadrado, obteniendo que la probabilidad de obtener un valor 7,81 o mayor con 3 grado de libertad es $p = 0,184$. Por tanto, el valor empírico obtenido en el ejemplo no sería estadísticamente significativo, pues es mayor que 0,01. La conclusión es que no se puede rechazar la hipótesis de homogeneidad de grupos sanguíneos en las dos muestras. Por tanto provisionalmente se asumiría la homogeneidad, aunque se sabe que la aceptación de la hipótesis no es tan fuerte como su rechazo (Messick, 1992).

Terminado el ejercicio, se recordaría la interpretación y el cálculo del p valor, que supuestamente ya conocerían los estudiantes, de forma general y que debieran aplicar en este caso.

El tema se cerraría dando las condiciones de aplicación de Chi-cuadrado, y algunas observaciones de interés (Martín y Luna, 1994):

- El valor de Chi-cuadrado en la tabla de la distribución, obtenemos siempre un valor positivo, por lo que, siempre hacemos un contraste unilateral.

- Si las frecuencias esperadas son muy pequeñas, entonces se obtendría un valor alto de Chi-cuadrado. Por eso, se recomienda que se use una muestra de suficiente tamaño. Dos recomendaciones importantes son las siguientes: (1) como máximo el 20% de las frecuencias esperadas pueden ser menores que el valor 5; (2) no debe usarse este contraste si hay frecuencias esperadas inferiores a 1.

A continuación se introduciría y se resolvería con los estudiantes el Ejercicio 3.1, adaptado de Guàrdia, Freixa, Però, y Turbany (2007).

Ejercicio 3.1. Inmigración.

Se desea estudiar hasta qué punto existe relación entre el tiempo de residencia de inmigrantes en nuestro país y su percepción de integración. Se dispone de una muestra pequeña de 230 inmigrantes a los que se les evaluó en ambas variables obteniéndose la siguiente tabla de frecuencias observadas. ¿Confirman estos datos la hipótesis planteada con un nivel de confianza del 95%?

Tiempo de Residencia	Grado de integración		Total
	Bajo	Alto	
Más tiempo	40	90	130
Menos tiempo	90	10	100
Total	130	100	230

(Adaptado de Guàrdia, Freixa, Però, y Turbany, 2007, p. 130).

Se calcularían las frecuencias esperadas, comprobando que ninguna es menor que 5, por lo que el método es aplicable y se calcularía el estadístico Chi-cuadrado, así como sus grados de libertad.

$$(n-1) \times (m-1) = 1 \times 1 = 1.$$

Tabla 4.5.5. Frecuencias esperadas

Tiempo de Residencia	Grado de integración	
	Bajo	Alto
Más tiempo	73,478	56,52
Menos tiempo	56,522	43,48

$$\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 80,69$$

Al buscar en la tabla de Chi-cuadrado, se obtendría un valor $p = 0,000$ para este Chi-cuadrado. Por tanto el Chi-cuadrado sería estadísticamente significativo y se rechazaría la hipótesis de independencia de las variables.

4.5.3. SIGNIFICADO INSTITUCIONAL PRETENDIDO

Para sintetizar el análisis de la lección 3 presentamos la Tabla 4.5.6. Observamos que en esta lección se presentan dos campos de problemas nuevos cuya solución requiere el estadístico Chi-cuadrado, y ambos relacionados con la inferencia estadística: (a) Estudiar la posible asociación o independencia de dos variables; y (b) Analizar la homogeneidad de varias muestras. Ambos campos de problemas se introducen desde una perspectiva inferencial, que conlleva la realización de un contraste de hipótesis.

Tabla 4.5.6. Síntesis del significado presentado en la tercera lección

Tipo de Objetos	Objetos matemáticos
Situaciones-problemas	P3. Contraste de homogeneidad.
	P4. Contraste de independencia.
Lenguaje	R5. Lenguaje simbólico.
	R7. Tablas (de frecuencia y de contingencia).
Procedimientos	PR1. Lectura e interpretación de la tabla.
	PR6. Cálculo de las frecuencias esperadas.
	PR7. Comprobar las propiedades de independencia.
	PR12. Formulación de hipótesis (contraste de independencia).
	PR13. Formulación de hipótesis (contraste de homogeneidad).
	PR14. Interpretación de resultados del test Chi-cuadrado obtenido con el ordenador en el contexto del problema.
	PR15. Completar un contraste de hipótesis tomando una decisión.
	PR17. Cálculo de los grados de libertad para realizar un test Chi-cuadrado.
	PR18. Manejo de tablas de distribuciones de probabilidad.
PR19. Cálculo del estadístico de contraste.	
Conceptos	C1. Variable simple / doble valor.
	C2. Distribución bidimensional.
	C4. Tablas de contingencia, tablas rxc.
	C5. Frecuencia absoluta y relativa doble.
	C6. Frecuencia absoluta y relativa marginal; distribución marginal.
	C7. Frecuencia relativa condicional; distribución condicional.
	C10. Independencia /Asociación, dependencia funcional.
	C11. Asociación inversa, directa.
	C12. Intensidad de la asociación.
	C17. Frecuencias esperadas.
	C18. Estadístico Chi-cuadrado, grados de libertad, distribución.
	C22. Significación.
	C23. p valor.
	C24. Contraste de hipótesis. Hipótesis alternativa y nula.
	C25. Población y muestra. Estadístico de contraste.
C26. Región crítica y de aceptación.	
C27. Posibles errores en un contraste de hipótesis.	
Propiedades	PP7. Frecuencias esperadas en caso de independencia.
	PP8. Las hipótesis de un contraste son complementarias.
	PP11. Relación entre regiones de aceptación y rechazo y decisión en el contraste.
	PP12. Relación entre valor p en el contraste Chi-cuadrado y asociación.
	PP17. Diferenciar valores de Chi-cuadrado estadísticamente significativos.
	PP18. Posibles errores en un contraste.
PP19. Condiciones de aplicación del estadístico Chi-cuadrado.	
Argumentos	A1. Razonamientos informales.
	A2. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar.

Se introducirían los nuevos lenguajes, verbales y simbólicos asociados a estos contrastes, por ejemplo, las gráficas y tablas de la distribución Chi-cuadrado, así como la notación habitual en el planteamiento de hipótesis. Se introducirían conceptos, procedimientos y propiedades correspondientes, recordando también los necesarios para llevar a cabo un contraste de hipótesis, como los de hipótesis nula y alternativa, región crítica y de aceptación o tipos de error en un contraste. Se recordaría la distribución Chi-cuadrado, su parámetro y propiedades, así como procedimientos asociados, tales como planteamiento de hipótesis, lectura de tablas, cálculo de grados de libertad, interpretación de resultados de un contraste y toma de decisión.

Como en las lecciones anteriores el profesor usaría argumentos informales para presentar el tema; también serían informales los esperados por parte de los estudiantes. La cadena de pasos para llevar a caso el contraste de hipótesis sería un razonamiento deductivo.

4.6. ANÁLISIS DE LA CUARTA LECCIÓN. MEDIDAS DE ASOCIACIÓN

4.6.1. OBJETIVOS Y CONTENIDOS

Se plantean como objetivos que los estudiantes aprendiesen a:

- Interpretar la intensidad de la dependencia entre dos variables en una tabla de contingencia, desde independencia perfecta a dependencia funcional, aceptando la existencia de diferente intensidad en la asociación estadística.
- Calcular e interpretar medidas de asociación en tablas 2x2: Coeficiente Phi de Pearson, Riesgo relativo y Razón de productos cruzados, en función del máximo posible que se puede alcanzar en cada una de estas medidas, o en términos probabilísticos, cuando no existe tal máximo. Para este tipo de tablas, interpretar el significado del signo de la asociación.
- Calcular e interpretar medidas de asociación en tablas $r \times c$, en función del máximo valor que pueden tomar en una tabla dada: Coeficiente de contingencia de Pearson y V de Cramer.
- Calcular e interpretar medidas de asociación que informan de la reducción del error de predicción de una variable, cuando se conoce el valor de la otra: Lambda de Goodman y Kruskal. Diferenciar entre variable dependiente e independiente en tales casos y reconocer la existencia de dos valores para estos coeficientes.

- Elegir la medida de asociación más adecuada, en función del problema propuesto.

Contenidos

En la cuarta lección, se introducirían las medidas de asociación, como coeficientes que permiten evaluar la mayor o menor intensidad de la misma, una vez aceptada la existencia de asociación entre las variables. Se diferenciarían las medidas de asociación para tablas 2x2 y rxc, introduciendo sus posibles valores e interpretación.

Sus contenidos se describen con detalle en el análisis del desarrollo (Sección 4.6.2) y se presentan organizados respecto al tipo de objeto matemático considerados en nuestro marco teórico en la Tabla 4.6.1.

4.6.2. DESARROLLO

Se comenzaría este tema recordando lo visto en la lección anterior, para de este modo establecer un vínculo entre los temas 3 y 4. Para ello se recordaría que el contraste Chi-cuadrado nos informa de la presencia o no de independencia entre dos variables estadística, pero una vez aceptada la existencia de asociación, sería de interés poder medir su intensidad. Se indicaría que en este tema se les enseñaría las diferentes medidas existentes para saber la fuerza de la asociación, distinguiendo medidas para tablas 2x2 y para tablas rxc. Para cada medida se explicaría su cálculo, el rango en el que se mueve y su interpretación.

Comenzaríamos por explicar las medidas más habituales para estudiar la asociación en tablas 2x2: coeficiente Phi de Pearson, riesgo relativo (por filas y por columnas) y razón de productos cruzados (Ato y López, 1996). Se les mostraría los pasos requeridos en el cálculo de cada uno de estos estadísticos en el ejemplo 4.1, sobre observación de conducta, tomado de Solanas, Salafranca, Fauquet y Núñez (2005).

Ejemplo 4.1. Observación de conducta

Para realizar un estudio de observación de conductas de interacción en niños en situación de juego se ha entrenado a dos observadores en la utilización de un sistema de registro de conductas. Los dos observadores codifican con el mismo sistema de categorías, requiriéndose que lo utilicen con un mismo criterio. Para evaluar el nivel de acuerdo entre los observadores y constatar si el entrenamiento recibido ha sido adecuado, se pide a ambos observadores que clasifiquen las conductas observadas en un vídeo de prueba. Los resultados fueron los siguientes:

Observador O ₁	Observador O ₂		Total
	A	B	
A	100	10	110
B	20	60	80
Total	120	70	190

Calcula e interpreta el coeficiente Phi de Pearson como valor de la intensidad.
(Adaptado de Solanas, Salafranca, Fauquet y Núñez, 2005, p. 424)

Como primer paso en este ejemplo, se calcularía el valor Chi-cuadrado:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 86,47$$

Se razonaría que, al ser una suma de cuadrados, el valor de Chi-cuadrado es mayor conforme aumenta el número de sumandos, por lo que su valor absoluto no está acotado, lo que lo hace difícil de interpretar como medida de asociación. Se sugeriría la necesidad de obtener un coeficiente que no dependiese de este tamaño, motivando la definición del coeficiente Phi de Pearson, que en este ejemplo tomaría el siguiente valor:

$$\Phi = \sqrt{\chi^2 / n} = \sqrt{86,47 / 190} = 0,675$$

Se observaría que el valor es siempre positivo, pero que, para determinar su signo, habría que estudiar las celdas de la diagonal principal, que en este caso, corresponden a la dependencia directa. En efecto, aparecerían muchos más datos en la diagonal principal f_{11} y f_{22} que en la otra diagonal. Se razonaría a los estudiantes que el valor es moderado-alto (cerca de 0,7), por lo que puede considerarse una asociación apreciable. Se seguiría desarrollando el ejemplo, como sigue:

Ejemplo 4.1. Observación de conducta (continuación)

Calculemos estos dos riesgos en el ejemplo sobre observación de conducta

El riesgo relativo por columnas sería en este ejemplo el siguiente:

$$RR_{\text{columnas}} = \frac{P(O_1_A / O_2_A)}{P(O_1_A / O_2_B)} = \frac{100/120}{10/70} = \frac{100 \times 70}{10 \times 120} = \frac{7000}{1200} = 5,8333$$

Al obtener un valor $RR_{\text{columnas}} > 1$, nos indicaría que existe asociación positiva (que antes se había confirmado). El significado de su valor se interpretaría diciendo que sería 5,8333 veces más fácil tener un valor A por el observador O_1 , cuando el observador O_2 también observa la conducta A que cuando el observador O_2 observa la conducta B.

El riesgo relativo por filas es:

$$RR_{\text{filas}} = \frac{P(O_2_A / O_1_A)}{P(O_2_A / O_1_B)} = \frac{100/110}{20/80} = \frac{100 \times 80}{20 \times 110} = \frac{8000}{2200} = 3,6364$$

El $RR_{filas} > 1$, nos indica que existe asociación positiva. Sería 3,6364 veces más fácil tener un valor A por el observador O_2 , cuando el observador O_1 también obtiene un valor A, que si este último obtiene un valor B. A continuación se continuaría, con otra parte del ejemplo:

Ejemplo 4.1. Observación de conducta (continuación)

Calculamos la razón de productos cruzado para este ejemplo

$$\text{Se obtendría: } RC = \frac{f_{11}f_{22}}{f_{12}f_{21}} = \frac{100 \times 60}{10 \times 20} = \frac{6000}{200} = 30$$

Al obtener un valor mayor que 1, entonces habría una dependencia directa. Entre los sujetos calificados como A por el observador O_2 , habría 100 valorados con A por el observador O_1 , por cada 20 valorados con B (la razón sería 5/1).

Se continuaría por las medidas de asociación en tablas rxc, siguiendo a Ruiz-Maya, Martín, Montero y Uriz (1995); dentro de ellas se les explicaría el coeficiente de contingencia de Pearson y el valor V de Cramer. Se les mostraría los pasos de cada uno en el ejemplo 4.2, sobre el aborto, tomado de Pardo y San Martín (1998):

Ejemplo 4.2. Aborto

Para analizar si el estado civil no era una variable relevante a la hora de explicar las actitudes abortistas, se ha encuestado a 500 sujetos obteniendo los resultados que aparecen en la tabla siguiente.

	Actitud Abortista	Actitud Antiabortista	Total
Solteros	120	30	150
Casados	50	200	250
Divorciados	30	70	100
Total	200	300	500

Calcula e interpreta el coeficiente de contingencia de Pearson y su máximo.
(Pardo y San Martín, 1998, p. 555)

Para este ejemplo, se pediría a los estudiantes calcular el valor Chi-cuadrado, para el cual se obtendría un valor estadísticamente significativo, lo que se podría comprobar leyendo la tabla de la distribución para 2 grados de libertad:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 145,83$$

Como para el caso de tablas 2x2, se razonaría que el Chi-cuadrado no permite valorar la intensidad de la asociación, motivando la definición del coeficiente C , que, para el ejemplo, se calcularía mediante la siguiente formula:

$$C = \sqrt{\chi^2 / (\chi^2 + n)} = \sqrt{145,83 / (145,83 + 500)} = 0,475$$

Para interpretar el valor obtenido, se indicaría a los estudiantes que el máximo posible para esta tabla sería:

$$Max \{C\} = \sqrt{\frac{Min\{r-1, c-1\}}{1 + Min\{r-1, c-1\}}} = 0,7071$$

Por ello, se razonaría que la asociación es moderada-alta, pues se acerca al 67% del máximo, con lo que la interpretación sería semejante a la de un coeficiente de correlación de 0,67.

Ejemplo 4.2. Aborto (continuación)
Calcula e interpreta el valor V de Cramer.

Se continuaría el ejemplo, calculando el coeficiente V mediante la siguiente formula, siendo $p = \text{Min}\{\text{número de filas, número de columnas}\}$:

$$V = \sqrt{\chi^2 / n(p-1)} = 0,54$$

Para interpretar el valor obtenido, se indicaría a los estudiantes que el máximo posible para este coeficiente es el valor 1 y se argumentaría que el valor obtenido indica que la asociación es moderada-alta, ya que es un poco superior al valor medio posible.

Se continuaría por las medidas basadas en la reducción proporcional del error en tablas $r \times c$ (Batanero y Díaz, 2008), motivando su introducción por el interés que tiene en algunos estudios diferenciar una de las variables como dependiente y otra como independiente. Se estudiarían los coeficientes Lambda de Goodman y Kruskal, mostrando su cálculo e interpretación en el ejemplo 4.2, sobre el aborto.

Ejemplo 4.2. Aborto (continuación)
Calcula e interpreta el valor Lambda de Goodman y Kruskal, para filas dependientes.

El coeficiente Lambda de Goodman y Kruskal, tomando las filas como variable dependiente, se calcularía mediante la siguiente formula:

$$\lambda = \frac{\sum f_{mj} - f_{\max}}{n - f_{\max}} = \frac{(120 + 200) - 250}{500 - 250} = 0,28$$

Siendo: f_{m+} la mayor frecuencia marginal en filas y f_{mj} la mayor frecuencia en la columna j -ésima. El error de predicción es la cuantificación de la reducción del error que se comete al predecir el valor de una variable, cuando se conoce el valor de la otra.

Se diría que la interpretación del valor del coeficiente supone un 28% del error al predecir el valor del tipo de actitud abortista, cuando se conoce el estado civil, en relación al cometido si no se tiene esta información; la mejora de la predicción no es muy buena, ya que el valor del coeficiente es bajo.

A continuación se resolvería el Ejercicio 4.1, con la participación de los estudiantes. Está tomado de Cuadras, Echeverría, Mateo y Sánchez (1984) y trata sobre calificaciones.

Ejercicio 4.1. Calificaciones de estadística.

Al final del curso los cuatro profesores de un mismo curso de estadística calificaron a los alumnos de la siguiente forma:

Profesores	NOTAS		Total
	Aprobados	Suspensos	
A	68	57	125
B	90	60	150
C	70	30	100
D	120	30	150
Total	348	177	525

a. Calcula los coeficientes de contingencia de Pearson V de Cramer y Lambda de Goodman y Kruskal (con filas dependientes)

b. ¿Qué se puede concluir de lo obtenido?

(Cuadras, Echeverría, Mateo y Sánchez, 1984, p. 637).

Se calcularía el valor Chi-cuadrado, que sería: $\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 23,80$.

De ello se esperaría que los estudiantes dedujesen que el contraste es estadísticamente significativo y hay asociación entre profesor y aprobados/suspensos.

- El coeficiente de contingencia de Pearson, para este ejemplo es $C = \sqrt{\chi^2 / (\chi^2 + n)} = 0,208$, con un máximo: $\text{Max } \{C\} = \sqrt{\frac{\text{Min}\{r-1, c-1\}}{1 + \text{Min}\{r-1, c-1\}}} = 0,7071$. Por tanto, los estudiantes descubrirían que la intensidad de la asociación es pequeña, pues sólo llegaríamos a la tercera parte del máximo posible.
- El coeficiente V de Cramer sería $V = \sqrt{\chi^2 / n(p-1)} = 0,213$. Como los estudiantes ya sabrían que este coeficiente puede llegar a 1, debieran confirmar que la asociación es pequeña.

- La Lambda de Goodman y Kruskal con filas dependientes

sería: $\lambda_x = \frac{(\sum f_{mj}) - f_{m+}}{N - f_{m+}} = 0,08$, de lo que los estudiantes observarían que no

mejora mucho la predicción al saber si el estudiante ha aprobado para predecir su profesor.

Se haría una discusión en la clase para concluir que todos los coeficientes estarían por debajo de la mitad del rango que pueden tomar, sin llegar a ser nulos y por tanto, se podría decir que resulta una asociación baja.

A continuación se pediría a los estudiantes que leyesen en los apuntes la solución del Ejercicio 4.2 (San Martín y Pardo, 1989), paso a paso, promoviendo un aprendizaje por descubrimiento. El profesor les ayudaría en las posibles dudas.

Ejercicio 4.2. Niños disléxicos

Un psicólogo escolar está probando la eficacia de tres métodos de lectura de niños disléxicos. Estudiando los antecedentes de estos niños vio que la experiencia en la familia de hermanos disléxicos podría ser una causa de la ineficacia de los métodos. Con el fin de obtener alguna evidencia, midió a sus alumnos en las dos variables y obtuvo la tabla siguiente de curación.

	Pacientes curados			Total
	Método A	Método B	Método C	
Sin hermanos disléxicos	0	6	14	20
Con hermanos disléxicos	10	16	4	30
Total	10	22	18	50

- Calcula el coeficiente de contingencia de Pearson.
- Calcula el coeficiente V de Cramer.
- Calcula la lambda de Goodman y Kruskal.
- ¿Qué se puede concluir de lo obtenido?

(Adaptado de San Martín, y Pardo, 1989, p. 444).

Para este ejercicio, el valor Chi-cuadrado sería: $\chi^2_{exp} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 18.86$;

el coeficiente de contingencia de Pearson $C = \sqrt{\chi^2 / (\chi^2 + n)} = 0,523$, con un máximo

$Max \{C\} = \sqrt{\frac{Min\{r-1, c-1\}}{1 + Min\{r-1, c-1\}}} = 0,707$; el coeficiente V de Cramer $V = \sqrt{\chi^2 / n(p-1)} =$

0,614 y la lambda de Goodman y Kruskal (filas dependientes): $\lambda_x = \frac{(\sum f_{mj}) - f_{m+}}{N - f_{m+}} =$

0,5.

Se espera que los estudiantes concluyesen que, al estar todos los coeficientes por encima de la mitad del rango que pueden tomar, sin llegar a ser el máximo se podría decir una asociación moderada-alta. Si hubiese tiempo se podrían comparar los dos ejemplos analizados y la intensidad de la asociación en cada caso.

4.6.3. SIGNIFICADO INSTITUCIONAL PRETENDIDO

Para sintetizar el análisis del Tema 4, se presenta un resumen del significado pretendido (Tabla 4.6.1).

Tabla 4.6.1. Síntesis del significado presentado en la cuarta lección

Tipo de objetos	Objetos matemáticos
Situaciones-problemas	P5. Definir una medida de la intensidad de asociación entre las variables
Lenguaje	R4. Lenguaje verbal. R5. Lenguaje simbólico. R7. Tablas (de frecuencia y de contingencia).
Procedimientos	PR1. Lectura e interpretación de la tabla.
	PR6. Cálculo de las frecuencias esperadas.
	PR7. Comprobar las propiedades de independencia.
	PR8. Selección de medidas de la asociación.
	PR9. Cálculo e interpretación de medidas en tablas 2x2.
	PR10. Cálculo e interpretación de medidas en tablas rxc.
	PR11. Cálculo e interpretación de medidas basadas en la reducción del error.
	PR14. Interpretación de resultados del test Chi-cuadrado obtenido con el ordenador en el contexto del problema PR19. Cálculo del estadístico de contraste.
Conceptos	C1. Variable simple / doble valor.
	C2. Distribución bidimensional
	C11. Asociación inversa, directa.
	C12. Intensidad de la asociación.
	C17. Frecuencias esperadas.
	C18. Estadístico Chi-cuadrado, grados de libertad, distribución.
	C19. Coeficiente Phi, Riesgo relativo, Razón de productos cruzados (medidas tablas 2x2).
	C20. Coeficiente C y V (medidas simétricas, tablas rxc).
	C21. Coeficiente Lambda (medidas asimétricas).
Propiedades	PP13. Diferencias entre el significado y aplicación de las medidas de asociación.
	PP19. Relación entre el riesgo relativo por filas y columnas.
	PP20. Rango de acción de las medidas de asociación.
	PP21. Relación entre probabilidad y tabla de contingencia.
Argumentos	A1. Razonamientos informales.
	A2. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar.
	A3. Deducción de propiedades sin excesiva formalización

Observamos que en esta lección se presentan el campo de problemas P5 consistente en definir medidas de asociación, para el cuál hemos considerado tres subcampos: (a) cálculo de medidas de asociación en tablas 2x2, (b) medidas de asociación en tablas rxc, y (c) medidas de asociación que informan de la reducción del

error de predicción de una variable. Estos coeficientes son necesarios para poder juzgar la intensidad de la asociación, una vez aceptada su existencia mediante un contraste de asociación.

A lo largo de este tema se introducirían nuevos lenguajes, verbales y simbólicos asociados a las medidas de asociación. Los principales conceptos serían dichas medidas; y se estudiarían las propiedades asociadas; en particular las que relacionan los valores posibles de los coeficientes con la intensidad y signo de la asociación. Como procedimientos, se destaca el cálculo de los coeficientes y la interpretación, así como la de su valor máximo.

Los argumentos usados por el profesor para presentar el tema, así como los esperados por parte de los estudiantes serían informales, comprobación de ejemplos y contraejemplos, así, como, en algún caso, razonamientos deductivos sencillos.

4.7. SESIÓN DE PRÁCTICAS

Además de las sesiones anteriormente analizadas, se planificaron dos sesiones prácticas de trabajo con el ordenador en la sala de informática. En la primera de dichas sesiones se comentaría como sería el trabajo en las clases de prácticas, el tipo ejercicios con los que se trabajaría y que se había preparado un programa.

El profesor informaría que la asistencia se controlaría, además de recordar que en la primera práctica se les ayudaría a manejar el programa y a resolver los problemas, pero que en la segunda debieran intentar resolverlos en forma individual y sin ayuda. Se les indicaría que no penalizan los errores y se premiaría el esfuerzo. Se les comunicaría que el material necesario (guión del programa y ejercicios) estaría disponible en la página web, y que las soluciones de las prácticas se colgarían también en dicha página una vez recogidas sus soluciones escritas. A continuación se iría introduciendo el funcionamiento y objetivos del programa, cuya versión completa se incluye en el Anexo 6.

4.7.1. PROGRAMA DE CÁLCULO EXCEL

Se explicaría que el programa Excel permite la utilización de varias ventanas, que facilitan los cálculos y resultados y está organizado en cinco hojas: “Frecuencias”, “Gráficos”, “Test Chi-cuadrado”, “Medidas asociación tablas 2x2” y “Medidas asociación tablas rxc”. A continuación se analiza su contenido.

Frecuencias

Este programa (Figura 4.7.1.1), colocado en la primera pestaña, permite resumir datos sobre las dos variables estadística que queramos analizar en una tabla de contingencia. Se comenzaría explicando los objetivos del programa que serían los siguientes.

1. Reconocer las dos variables estadísticas presentes en una tabla de contingencia. Se usaría como ejemplo unos datos tomados en Internet sobre supervivencia en el Titanic.
2. Diferenciar los valores que pueden tomar las dos variables estadísticas.
3. Identificar las frecuencias dobles que corresponde a cada par de valores (x_i, y_j) de las variables, que representamos mediante f_{ij} . Identificar el total de la muestra. En nuestro caso el valor es 1313 (número de pasajeros del Titanic).
4. Calcular las frecuencias relativas dobles y frecuencias marginales e interpretarlas.
5. Obtener las frecuencias condicionales e interpretarlas.

Se explicaría el significado de los cuadros sombreados en gris y verde, indicando que estas son las celdas que tenemos que rellenar nosotros y el mismo color se usa para señalar los datos que se pueden variar en todas las pestañas del programa. A continuación se explicarían los datos requeridos que son los siguientes:

- *Nombre de la variable X*, que varía de un ejemplo a otro, por lo que requiere su introducción, en nuestro ejemplo es “Clase”.
- *Nombre de la variable Y*. También cambia según el ejemplo; en nuestro ejemplo es “Supervivencia”.
- *Valores que pueden tomar las dos variables*. En el ejemplo del Titanic la variable X toma los valores $(x_1=Primera, x_2=Segunda$ y $x_3=Tercera)$. La variable Y toma los valores $(y_1=Si, y_2=No)$.
- *Frecuencias absolutas dobles*. Serían los valores de las celdas de la tabla; en el ejemplo considerado es una tabla 2x3.

Figura 4.7.1.1. Hoja “Frecuencias”

FRECUENCIAS DOBLES							FRECUENCIAS RELATIVAS POR FILAS							
Variable Y							Variable Y							
Supervivencia							Supervivencia							
y1 y2 y3 y4 y5							y1 y2 y3 y4 y5							
Si No							Si No 0 0 0							
Variable X	x1 Primera	194	128			Total	Variable X	x1 Primera	0,60	0,40	0,00	0,00	0,00	1
Clase	x2 Segunda	119	161			280	Clase	x2 Segunda	0,43	0,58	0,00	0,00	0,00	1
	x3 Tercera	138	573			711		x3 Tercera	0,19	0,81	0,00	0,00	0,00	1
	x4					0		x4	0,00	0,00	0,00	0,00	0,00	0
	x5					0		x5	0,00	0,00	0,00	0,00	0,00	0
	Total	451	862	0	0	1313								
FRECUENCIAS RELATIVAS DOBLES							FRECUENCIAS RELATIVAS POR COLUMNAS							
Variable Y							Variable Y							
Supervivencia							Supervivencia							
y1 y2 y3 y4 y5							y1 y2 y3 y4 y5							
Si No 0 0 0							Si No 0 0 0							
Variable X	x1 Primera	0,15	0,10	0,00	0,00	0,00	0,25	Variable X	x1 Primera	0,43	0,15	0,00	0,00	0,00
Clase	x2 Segunda	0,09	0,12	0,00	0,00	0,00	0,21	Clase	x2 Segunda	0,26	0,19	0,00	0,00	0,00
	x3 Tercera	0,11	0,44	0,00	0,00	0,00	0,54		x3 Tercera	0,31	0,66	0,00	0,00	0,00
	x4	0	0,00	0,00	0,00	0,00	0,00		x4	0	0,00	0,00	0,00	0,00
	x5	0	0,00	0,00	0,00	0,00	0,00		x5	0	0,00	0,00	0,00	0,00
	Total	0,34	0,66	0,00	0,00	0,00	1,00	Total	1,00	1,00	0,00	0,00	0,00	

NOTA IMPORTANTE. SÓLO PUEDES ESCRIBIR EN LOS CUADROS VERDES

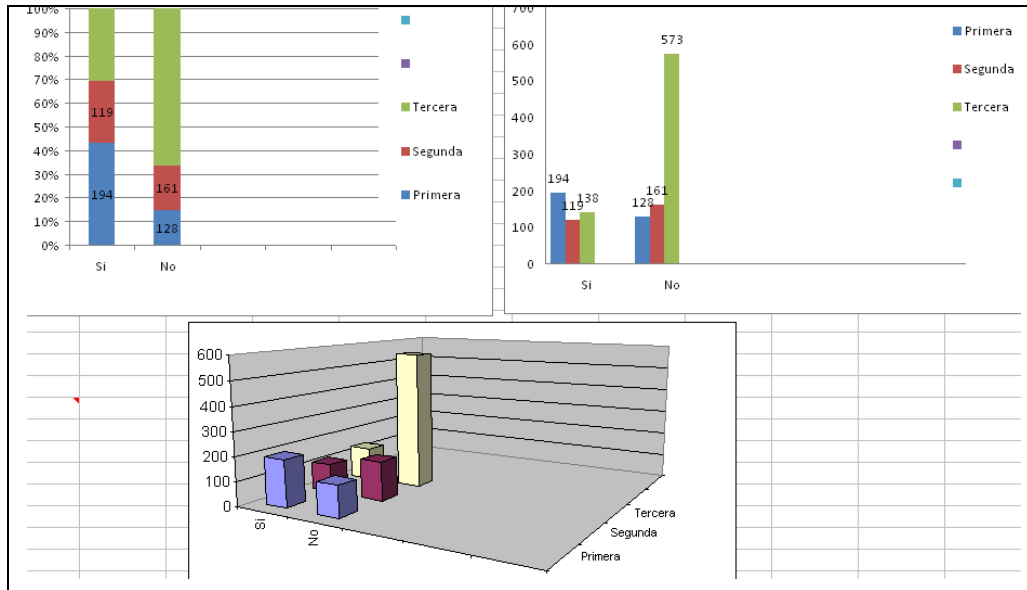
Finalmente se mostrarían los resultados del programa que son los siguientes, recordando su definición e interpretación: Frecuencias marginales, por filas y columnas; número total de sujetos en el experimento; tabla con las frecuencias relativas dobles, con sus totales; tabla con las frecuencias relativas por columnas, con sus totales y tabla con las frecuencias relativas por filas, con sus totales.

Gráficos

La segunda pestaña (Figura 4.7.1.2), nos aporta información visual. Está constituida por los gráficos relacionados con las tablas de contingencia que los alumnos habían estudiado en el Tema 1, y la mayoría de los cuales ya conocían de cursos anteriores. Para explicar esta pestaña, se comenzaría con los objetivos del programa que serían los siguientes.

1. Observar e interpretar diagrama de barras adosadas correspondientes a una tabla de contingencia, analizando el tipo de información que proporcionan y el cambio en la representación al intercambiar filas y columnas.
2. Observar e interpretar los datos mediante diagrama de barras apiladas, observando como éste permite visualizar mejor la asociación en la tabla de contingencia que el diagrama adosado; también se intercambiarían filas y columnas.
3. Observar e interpretar los datos mediante gráfico tridimensional, comparando las tres representaciones entre sí

Figura 4.7.1.2. Hoja "Gráficos"



Se explicaría el procedimiento para girar el gráfico tridimensional, de modo que se tenga una mejor visualización del mismo. Se recordaría que en esta pestaña no hay celdas en verde, por lo que no habría que completar ningún dato en esta parte. Finalmente se mostrarían los resultados del programa que son la representación gráfica de los datos mediante diagrama de barras adosadas, diagrama de barras apiladas y gráfico tridimensional.

Test Chi-cuadrado

La tercera pestaña (Figura 4.7.1.3), permite el cálculo del estadístico Chi-cuadrado, a partir de las tablas de contingencia, para poder utilizarlo tanto en los contrastes de homogeneidad como en los de asociación. Se comenzaría explicando los objetivos del programa que serían los siguientes.

1. Realizar e interpretar el cálculo de las frecuencias esperadas en caso de independencia a partir de una tabla de contingencia.
2. Cálculo e interpretación de las componentes de Chi-cuadrado.
3. Cálculo e interpretación del estadístico Chi-cuadrado y de sus grados de libertad.
4. Interpretación del valor Chi-cuadrado y del valor p .

Se recordaría el significado de los cuadros sombreados en gris y verde, explicando que los datos requeridos en esta parte son únicamente son los grados de libertad. Finalmente se mostrarían los resultados del programa que son los siguientes:

- Tabla de frecuencias absolutas dobles, con las frecuencias marginales su total. Se repite en esta pestaña las frecuencias observadas para un estudio más cómodo por parte de los alumnos
- Tabla de frecuencias esperadas en caso de independencia.
- Tabla con las componentes de Chi-cuadrado, estadístico Chi-cuadrado y valor p .

Figura 4.7.1.3. Hoja “Test Chi-cuadrado”

FRECUCIAS OBSERVADAS								TEST CHI CUADRADO							
								COMPONENTES DE CHI CUADRADO							
Variable Y								Variable Y							
Supervivencia								Supervivencia							
y1 y2 y3 y4 y5								y1 y2 y3 y4 y5							
Si No 0 0 0 Total								Si No 0 0 0 Total							
x1	Primera	194	128	0	0	0	322	Variable Xx1	Primera	62,883	32,9	0	0	0	95,783
x2	Segunda	119	161	0	0	0	280	Clase	x2	Segunda	5,4161	2,834	0	0	8,24982
x3	Tercera	138	573	0	0	0	711		x3	Tercera	46,199	24,17	0	0	70,3703
x4		0	0	0	0	0	0		x4		0	0	0	0	0
x5		0	0	0	0	0	0		x5		0	0	0	0	0
Total		451	862	0	0	0	1313		Total	114,5	59,91	0	0	0	174,40

FRECUCIAS ESPERADAS EN CASO DE INDEPENDENCIA							
Variable Y							
Supervivencia							
y1 y2 y3 y4 y5							
Si No 0 0 0 Total							
x1	Primera	110,6	211,4	0	0	0	322
x2	Segunda	96,177	183,8	0	0	0	280
x3	Tercera	244,22	466,8	0	0	0	711
x4		0	0	0	0	0	0
x5		0	0	0	0	0	0
Total		451	862	0	0	0	1313

Chi cuadrado=	174,40
Grados libertad=	2
p=	0,000

Medidas de asociación para tablas 2x2

La cuarta pestaña (Figura 4.7.1.4) sólo es utilizable en el caso de que nuestro ejercicio este constituido por una tabla de contingencia 2x2. En ella esta aparecen diferentes medidas de la intensidad de la asociación para este tipo de tablas. Se comenzaría explicando los objetivos del programa que serían los siguientes.

1. Calcular e interpretar el valor Phi de Pearson a partir de los datos de una tabla de contingencia 2x2.
2. Calcular e interpretar, el valor riesgo relativo (por filas y columnas). Diferenciar estas dos medidas y saber cuándo hay que aplicar una u otra.
3. Calcular e interpretar, el valor razón de productos cruzados y el valor del riesgo relativo por filas y por columnas, sabiendo cuando conviene calcular una de estas dos últimas medidas.

Se recordaría que en esta pestaña no hay celdas en verde, por lo que no habría que completar ningún dato y finalmente se mostrarían los resultados del programa en un ejemplo.

Figura 4.7.1.4. Hoja “Medidas de asociación 2x2”

TABLA DE DATOS					
		Variable Y			
		Supervivencia			
		y1	y2		
		Si	No	Total	
Variable X	x1 Primera	194	128	322	
Clase	x2 Segunda	119	161	280	
	Total	313	289	602	
				Chi cuadrado	174,40
				Phi de Pearson	0,538
				Riesgo Relativo	1,39941094 Columnas
					1,41761052 Filas
				Razón de productos cruzados	2,05055147

Medidas de asociación para tablas rxc

La quinta pestaña (Figura 4.7.1.5) sólo está pensada para usarse en el caso de que nuestro ejercicio este constituido por una tabla de contingencia rxc. En ella encontraremos diferentes medidas de la intensidad de la asociación para este tipo de tablas. Se comenzaría explicando los objetivos del programa que serían los siguientes.

1. Calcular e interpretar el valor coeficiente de contingencia de Pearson.
2. Calcular e interpretar el valor Lambda de Goodman y Kruskal.
3. Calcular e interpretar el valor V de Cramer.
4. Reconocer que el valor del coeficiente de contingencia de Pearson no alcanza un máximo fijo.

Se explicarían los datos requeridos que son el número de filas y columnas, mostrando finalmente los resultados del programa con un ejemplo.

Figura 4.7.1.5. Medidas de asociación rxc

		FRECUENCIAS DOBLES						
		Variable Y						
		Supervivencia						
		y1	y2	y3	y4	y5		
		Si	No	0	0	0	Total	
Variable X	x1 Primera	194	128	0	0	0	322	
Clase	x2 Segunda	119	161	0	0	0	280	
	x3 Tercera	138	573	0	0	0	711	
	x4 0	0	0	0	0	0	0	
	x5 0	0	0	0	0	0	0	
	Total	451	862	0	0	0	1313	
	Número de filas		3					
	Número de columnas		2					
	Coficiente de Contingencia de Pea			0,342	Máximo=		0,70710678	
	Lambda de Goodman y Kruskal			0,093				
	V de Cramer			0,364				

4.7.2. EJERCICIOS DE LA PRIMERA SESIÓN DE PRÁCTICAS

Como se ha dicho, en la primera práctica, el profesor ayudó a los alumnos a resolver los problemas, mientras que la segunda, que fue parte de la evaluación, los alumnos trabajaron solos. Por tanto, esta segunda parte se describirá en el Capítulo 6 (donde se analizan las respuestas a los ítems abiertos de evaluación). A continuación se analizan los cuatro problemas propuestos en la primera sesión práctica. Para los dos primeros no se requiere el uso del ordenador, aunque el estudiante podría emplearlo si lo desea.

Análisis del problema 1

Problema 1. En un colegio se pregunta a los alumnos, obteniendo los siguientes resultados:			
	Chicos	Chicas	Total
Le gusta el tenis	400	200	600
No le gusta	50	50	100
Total	450	250	700
Si elegimos al azar uno de estos alumnos:			
a.	¿Cuál es la probabilidad de que le guste el tenis?		
b.	¿Cuál es la probabilidad de que sea una chica y además le guste el tenis?		
c.	Sabiendo que el alumno elegido es chica, ¿Cuál es la probabilidad de que le guste el tenis?		

Este problema se puede resolver con o sin la ayuda del ordenador. En el primer apartado se pide el cálculo de una probabilidad simple, es decir $P(\text{guste el tenis})$, proporcionándole los datos mediante una tabla de contingencia 2x2. Para resolverlo, el estudiante ha de interpretar el enunciado, reconocer el significado de los valores de las

celdas y recordar la regla de Laplace. El estudiante ha de leer la tabla, identificar las variables (sexo y si le gusta ó no el tenis) y poblaciones en estudio. Una posible formulación correcta de la probabilidad sería la siguiente, donde se establece como numerador el número de alumnos que les gusta el tenis (frecuencia absoluta marginal), y como denominador el número de alumnos presentes (n).

$$P(\text{Gusta_Tenis}) = \frac{\text{Número_Alumnos_Gusta_Tenis}}{\text{Número_Alumnos}} = \frac{600}{700} = 0,857$$

En el segundo apartado se plantea el cálculo de una probabilidad conjunta, es decir $P(\text{ser chica y guste el tenis})$, en la misma tabla de contingencia 2x2. El alumno después de leer e interpretar el enunciado y la tabla, como en el apartado anterior, tiene que replantear el cálculo de está probabilidad de forma que pueda identificar el valor del numerador y del denominador. Una respuesta correcta sería la siguiente, donde se establece como numerador la frecuencia doble de alumnos que les gusta el tenis y son mujeres y como denominador el número total de alumnos (n).

$$P(\text{Gusta_Tenis} \cap \text{Chicas}) = \frac{\text{Número_Chicas_Gusta_Tenis}}{\text{Número_Alumnos}} = \frac{200}{700} = 0,2857$$

El tercer apartado pide al estudiante calcular una probabilidad condicional, $P(\text{guste el tenis} \mid \text{ser chica})$. El alumno comienza nuevamente con la lectura e interpretación, buscando el nuevo valor del numerador y del denominador. Una posible formulación correcta de esta probabilidad es establecer como numerador el número de alumnos que les gusta el tenis y son mujeres (frecuencia absoluta doble), y como denominador el número de chicas presentes (frecuencia absoluta marginal).

$$P(\text{Gusta_Tenis} \mid \text{Chicas}) = \frac{\text{Número_Chicas_Gusta_Tenis}}{\text{Número_Chicas}} = \frac{200}{250} = 0,8$$

En caso de resolver el problema con ayuda del ordenador, mediante la pestaña “frecuencias”, puede obtener las probabilidades pedidas (Ver Figura 4.7.2.1). La probabilidad simple no viene dada directamente. La probabilidad conjunta se obtiene de la tabla de frecuencias relativas dobles (redondeada a dos decimales). La probabilidad condicional, vendría incluida en la misma pestaña pero en la tabla que contiene las frecuencias condicionales por filas.

Figura 4.7.2.1. Solución del problema 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
3				Variable Y									Variable Y					
4				Sexo									Sexo					
5				y1	y2	y3	y4	y5					y1	y2	y3	y4	y5	
6				Chicos	Chicas				Total				Chicos	Chicas	0	0	0	Total
7	Variable X	x1 Si		400	200				600	Variable X	x1 Si		0,67	0,33	0,00	0,00	0,00	1
8	Gusta Tenis	x2 No		50	50				100	Gusta Tenis	x2 No		0,50	0,50	0,00	0,00	0,00	1
9		x3	0						0		x3	0	0,00	0,00	0,00	0,00	0,00	0
10		x4	0						0		x4	0	0,00	0,00	0,00	0,00	0,00	0
11		x5	0						0		x5	0	0,00	0,00	0,00	0,00	0,00	0
12		Total		450	250	0	0	0	700									
13																		
14				FRECUCIAS RELATIVAS DOBLES									FRECUCIAS RELATIVAS POR COLUMNAS					
15				Variable Y									Variable Y					
16				Sexo									Sexo					
17				y1	y2	y3	y4	y5					y1	y2	y3	y4	y5	
18				Chicos	Chicas	0	0	0	Total				Chicos	Chicas	0	0	0	Total
19	Variable X	x1 Si		0,57	0,29	0,00	0,00	0,00	0,86	Variable X	x1 Si		0,89	0,80	0,00	0,00	0,00	
20	Gusta Tenis	x2 No		0,07	0,07	0,00	0,00	0,00	0,14	Gusta Tenis	x2 No		0,11	0,20	0,00	0,00	0,00	
21		x3	0	0,00	0,00	0,00	0,00	0,00	0,00		x3	0	0,00	0,00	0,00	0,00	0,00	
22		x4	0	0,00	0,00	0,00	0,00	0,00	0,00		x4	0	0,00	0,00	0,00	0,00	0,00	
23		x5	0	0,00	0,00	0,00	0,00	0,00	0,00		x5	0	0,00	0,00	0,00	0,00	0,00	
24		Total		0,64	0,36	0,00	0,00	0,00	1,00		Total		1,00	1,00	0,00	0,00	0,00	
25																		
26																		
27	NOTA IMPORTANTE: SÓLO PUEDES ESCRIBIR EN LOS CUADROS VERDES																	

Análisis del problema 2

Problema 2. Un psicólogo en un estudio obtuvo la siguiente clasificación de los pacientes de un hospital de acuerdo con las variables “sexo” y “diagnóstico”.

Sexo	Diagnóstico	
	Esquizofr.	No esquizofr.
Varón	43	15
Mujer	32	52

- ¿Hay asociación entre las variables? ¿Es directa o inversa?
- ¿Por qué?

(Adaptado de Pardo y San Martín, 1998, p. 524)

Para resolver el problema el estudiante ha de hacer un juicio de asociación, a partir de los datos mediante una tabla de contingencia 2x2. Como en la primera sesión, todavía no se ha estudiado el contraste Chi cuadrado se espera que utilicen procedimientos más informales. El alumno podría intentar comparar los valores de las celdas que informan de una dependencia directa (43 y 52), con los valores de las celdas que informan de una dependencia inversa (15 y 32), llegando a la conclusión que hay más sujetos en el caso de dependencia directa. Sin embargo, sabemos que esta solución es sólo parcialmente correcta; por lo que sería más adecuado que el estudiante comparase las distribuciones frecuencias condicionales por filas o columnas, para comprobar si son iguales o diferentes.

El ordenador puede ayudarles en este procedimiento mediante la pestaña

“frecuencias”, donde se pueden observar las frecuencias relativas dobles (Figura 4.7.2.2). Debe ser capaz de diferenciar si necesita las frecuencias relativas condicionales por filas o columnas para resolver el problema.

Figura 4.7.2.2. Solución del problema 2

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
3				Variable Y									Variable Y					
4				Esquizofrénico									Esquizofrénico					
5				y1	y2	y3	y4	y5					y1	y2	y3	y4	y5	
6				Si	No				Total				Si	No	0	0	0	Total
7	Variable X	x1	Varón	43	15				58	Variable X	x1	Varón	0,74	0,26	0,00	0,00	0,00	1
8	Sexo	x2	Mujer	32	52				84	Sexo	x2	Mujer	0,38	0,62	0,00	0,00	0,00	1
9			x3						0		x3	0	0,00	0,00	0,00	0,00	0,00	0
10			x4						0		x4	0	0,00	0,00	0,00	0,00	0,00	0
11			x5						0		x5	0	0,00	0,00	0,00	0,00	0,00	0
12			Total	75	67	0	0	0	142									
13																		
14			FRECUCIAS RELATIVAS DOBLES									FRECUCIAS RELATIVAS POR COLUMNAS						
15				Variable Y									Variable Y					
16				Esquizofrénico									Esquizofrénico					
17				y1	y2	y3	y4	y5					y1	y2	y3	y4	y5	
18				Si	No	0	0	0	Total				Si	No	0	0	0	
19	Variable X	x1	Varón	0,30	0,11	0,00	0,00	0,00	0,41	Variable X	x1	Varón	0,57	0,22	0,00	0,00	0,00	
20	Sexo	x2	Mujer	0,23	0,37	0,00	0,00	0,00	0,59	Sexo	x2	Mujer	0,43	0,78	0,00	0,00	0,00	
21			x3	0	0,00	0,00	0,00	0,00	0,00		x3	0	0,00	0,00	0,00	0,00	0,00	
22			x4	0	0,00	0,00	0,00	0,00	0,00		x4	0	0,00	0,00	0,00	0,00	0,00	
23			x5	0	0,00	0,00	0,00	0,00	0,00		x5	0	0,00	0,00	0,00	0,00	0,00	
24			Total	0,53	0,47	0,00	0,00	0,00	1,00		Total		1,00	1,00	0,00	0,00	0,00	
25																		
26																		
27	NOTA IMPORTANTE: SÓLO PUEDES ESCRIBIR EN LOS CUADROS VERDES																	

Análisis del Problema 3.

Problema 3. Para comprobar si existía relación entre el tabaco y la hipertensión, fueron elegidas aleatoriamente 100 personas, de las que 55 resultaron ser hipertensas. Agrupadas dichas personas en tres clases, según que no fumaran, fumaran moderadamente o fumaran mucho, se encontró que entre los hipertensos 10 no fumaban y 25 lo hacían moderadamente. En cambio, entre los no hipertensos 18 fumaban moderadamente y 7 fumaban mucho.

- Organiza los datos para formar una tabla de doble entrada o de contingencia.
- Calcula las distribuciones condicionales por filas.
- Dibuja el diagrama de barras adosado y diagrama de barras apilado.

(Adaptado de Amón, 1993, pág. 386)

Para construir la tabla de contingencia en el primer apartado, el estudiante comenzaría con la lectura comprensiva del enunciado, identificando la población (100 personas), y las variables hipertensión (ser ó no hipertenso) y nivel de fumador (no fuma, fuma moderadamente ó fuma mucho). Posteriormente se tienen que identificar los valores del enunciado, para saber colocarlos en la tabla, considerando que 100 es el valor total (n); 55 es una frecuencia marginal absoluta de ser hipertenso; 10 es la frecuencia absoluta doble de no fumar y ser hipertenso; 25 es la frecuencia absoluta doble de fumar moderadamente y ser hipertenso; 18 es la frecuencia absoluta doble de fumar moderadamente y no ser hipertenso; y 7 es la frecuencia absoluta doble de fumar

mucho y no ser hipertenso. Con esta información plantearíamos la Tabla 4.7.2.1.

Tabla 4.7.2.1. Tabla de contingencia inicial

	Fumar			Total
	No fuma	Fuma moderadamente	Fuma mucho	
No ser hipertenso		18	7	
Ser hipertenso	10	25		55
Total				100

El alumno tiene que entablar relaciones entre las frecuencias absolutas dobles, las frecuencias absolutas marginales y el total de sujetos para completar los huecos que faltan en la tabla. El mejor método es buscar las filas y columnas que únicamente les falte un valor (incluido el valor del total). Por ejemplo, en la columna de la derecha correspondientes a las frecuencias marginales, tenemos: que $X + 55 = 100$; por tanto $X = 100 - 55 = 45$, y así sucesivamente hasta que queda la Tabla 4.7.2.2, que tendría que incluir en el ordenador.

Tabla 4.7.2.2. Tabla de contingencia final

	Fumar			Total
	No fuma	Fuma moderadamente	Fuma mucho	
No ser hipertenso	20	18	7	45
Ser hipertenso	10	25	20	55
Total	30	43	27	100

En el segundo apartado, el alumno, una vez construida la tabla de contingencia (apartado a), está en condiciones de hacer cálculos en ella e identificar en la hoja Excel la correspondiente a las filas (Tabla 4.7.2.3).

Tabla 4.7.2.3. Distribuciones condicionales por filas

		Variable Y					Total
		Fumar					
		y1	y2	y3	y4	y5	
		No	Moder.	Mucho	0	0	
Variable X	x1 No	0,44	0,40	0,16	0,00	0,00	1
Hipertensión	x2 Si	0,18	0,45	0,36	0,00	0,00	1

En el tercer apartado, el estudiante tiene que hacer dos diagramas que obtiene en forma automática mediante el programa Excel, (Figuras 4.7.2.1 y 4.7.2.2). Se discutiría con los estudiantes el hecho de que el segundo visualiza mejor la asociación entre las variables.

Figura 4.7.2.1. Diagrama de barras adosado

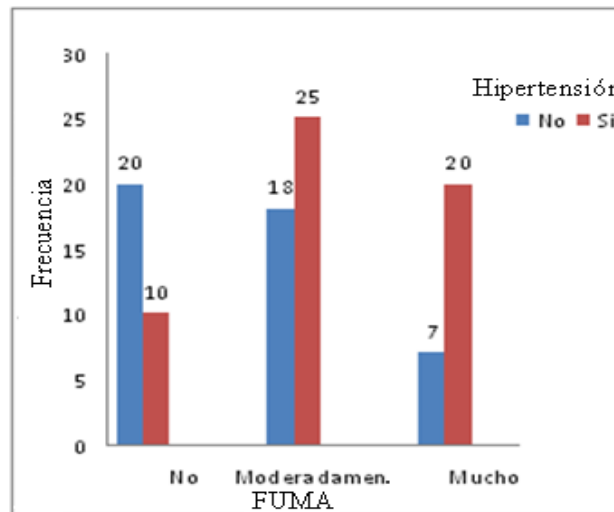
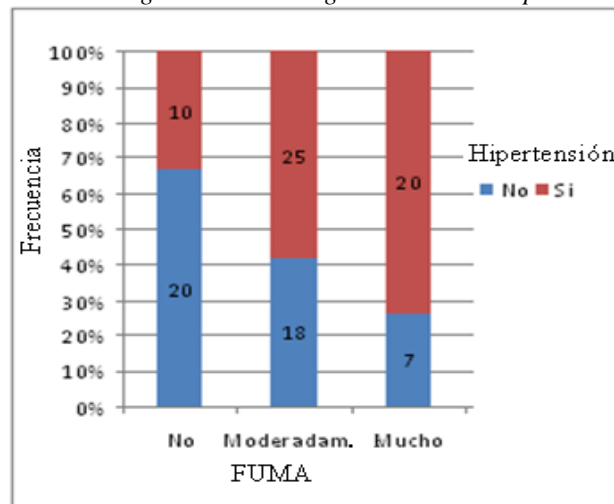


Figura 4.7.2.2. Diagrama de barras apilado



Análisis del problema 4

Problema 4. Un psicólogo en un estudio obtuvo la siguiente clasificación de una muestra de 176 sujetos de acuerdo con las variables, “concepción de la inteligencia” y “tipo de autoinstrucciones”

Concepción Inteligencia	Autoinstrucción		
	Respuestas Instruccionales	Respuestas Atribucionales	Otras respuestas
Destreza	24	11	9
Rasgo	72	33	27

Comprueba si se cumplen las propiedades de independencia.
(Adaptado de Pardo y San Martín, 1998, pág. 524)

Se espera que los estudiantes comprueben las principales propiedades de la independencia, para este ejemplo, utilizando los cálculos proporcionados por el programa Excel. Estas propiedades se analizan a continuación.

Propiedad 1. Las frecuencias absolutas dobles de los valores de la variable Y , son proporcionales en cada valor de la variable X . Es decir, las frecuencias relativas condicionales de los valores de la variable Y son iguales en cada valor de la variable X . El estudiante comenzaría con una lectura del enunciado y de la tabla, introduciéndola en el ordenador. Para la obtención de las frecuencias condicionales por filas nos fijaríamos en la pestaña “Frecuencias”, obteniendo la Tabla 4.7.2.4, en la que, efectivamente son idénticas en las dos filas. El mismo proceso lo puede repetir por columnas para ver que también en este caso coinciden.

Tabla 4.7.2.4. Distribuciones condicionales por filas

		Variable Y					
		Autoinstrucción					
		y1	y2	y3	y4	y5	
		Instruc.	Atrib.	Otras	0	0	Total
Variable X	x1 Destreza	0,55	0,25	0,20	0,00	0,00	1
Inteligencia	x2 Rasgo	0,55	0,25	0,20	0,00	0,00	1

Propiedad 2. La frecuencia relativa doble es igual al producto de las frecuencias relativas marginales de su fila y su columna: $h_{i,j} = h_{i.} \cdot h_{.j}$, para todo i, j . El estudiante podría comprobar la propiedad calculando manualmente y, comparando luego con las frecuencias relativas dobles y las frecuencias relativas marginales las obtienen en la pestaña “Frecuencias” como aparece en la Tabla 4.7.2.5.

Tabla 4.7.2.5. Frecuencias relativas dobles.

		Variable Y					
		Autoinstrucción					
		y1	y2	y3	y4	y5	
		Instruc.	Atrib.	Otras	0	0	Total
Variable X	x1 Destreza	0,14	0,06	0,05	0,00	0,00	0,25
Inteligencia	x2 Rasgo	0,41	0,19	0,15	0,00	0,00	0,75
	x3 0	0,00	0,00	0,00	0,00	0,00	0,00
	x4 0	0,00	0,00	0,00	0,00	0,00	0,00
	x5 0	0,00	0,00	0,00	0,00	0,00	0,00
	Total	0,55	0,25	0,20	0,00	0,00	1,00

$$\begin{aligned}
 h_{11} &= h_{1.} h_{.1} ; 0,14 = 0,25 \times 0,55 & h_{12} &= h_{1.} h_{.2} ; 0,06 = 0,25 \times 0,25 & h_{13} &= h_{1.} h_{.3} ; 0,05 = 0,25 \times 0,2 \\
 h_{21} &= h_{2.} h_{.1} ; 0,41 = 0,75 \times 0,55 & h_{22} &= h_{2.} h_{.2} ; 0,19 = 0,75 \times 0,25 & h_{23} &= h_{2.} h_{.3} ; 0,15 = 0,75 \times 0,2
 \end{aligned}$$

Propiedad 3. Si las variables son independientes, la frecuencia absoluta doble en cada celda de la tabla es igual a su frecuencia esperada. Para la comprobación de esta propiedad, el estudiante tiene que comparar las frecuencias absolutas, introducidas en la pestaña “Frecuencias”, con las frecuencias esperadas situadas en la pestaña “Test Chi cuadrado” (Tabla 4.7.2.6), viendo que coinciden estas dos frecuencias lo que implica independencia.

Tabla 4.7.2.6. Frecuencias esperadas en caso de independencia

		y1	y2	y3	y4	y5	
		Instruc.	Atrib.	Otras	0	0	Total
x1	Destreza	24	11	9	0	0	44
x2	Rasgo	72	33	27	0	0	132
x3	0	0	0	0	0	0	0
x4	0	0	0	0	0	0	0
x5	0	0	0	0	0	0	0
	Total	96	44	36	0	0	176

4.7.3. SIGNIFICADO INSTITUCIONAL PRETENDIDO

Como resumen del análisis de la sesión práctica que se ha llevado a cabo en las anteriores secciones, se presenta una síntesis del significado pretendido en la primera sesión de practicas (Tabla 4.7.3.1). Observamos que en esta práctica se presentan como campos de problemas todos los tratados en los dos primeros temas, al enseñar al estudiante a utilizar las diferentes hojas del programa.

Se introducirían los diversos elementos del lenguaje gráfico, recordando dichos gráficos si fuese necesario. También se requieren nuevos elementos de los lenguajes verbal y simbólico, y se recuerdan los conceptos, procedimientos y propiedades asociados a los temas teóricos. Los argumentos usados por el profesor para el funcionamiento de la práctica, así como los esperados por parte de los estudiantes serían informales, utilizándose también los ejemplos y contraejemplos. En algún caso de harán deducciones informales.

Tabla 4.7.3.1. Elementos de significado previstos en la sesión de prácticas

Tipo de objetos	Objetos matemáticos
Situaciones- problemas	P1. Resumir la información obtenida de un conjunto de observaciones bivariantes de modo que se visualicen posibles relaciones entre las variables.
	P2. Cálculo de probabilidades asociadas a valores específicos de una o las dos variables.
	P3. Contraste de homogeneidad
	P4. Contraste de independencia
	P5. Definir una medida de la intensidad de asociación entre las variables
Lenguaje	R1. Diagrama de barras adosado.
	R2. Diagrama de barras apilado.
	R3. Gráfico tridimensional.
	R4. Lenguaje verbal.
	R5. Lenguaje simbólico.
	R7. Tablas (de frecuencia y de contingencia).
Procedimientos	PR1. Lectura e interpretación de la tabla.
	PR2. Construcción de una tabla de contingencia.
	PR3. Cálculo de probabilidades simples, compuestas y condicionales a partir de la tabla.
	PR4. Cálculo de frecuencias condicionales y marginales (absolutas o relativas).
	PR5. Representación gráfica de datos a partir de una tabla de contingencia.
	PR6. Cálculo de las frecuencias esperadas.
	PR7. Comprobar las propiedades de independencia.
	PR8. Selección de medidas de la asociación.
	PR9. Cálculo e interpretación de medidas en tablas 2x2.
	PR10. Cálculo e interpretación de medidas en tablas rxc.
	PR11. Cálculo e interpretación de medidas basadas en la reducción del error.
	PR12. Formulación de hipótesis (contraste de independencia).
	PR13. Formulación de hipótesis (contraste de homogeneidad).
	PR14. Interpretación de resultados del test Chi-cuadrado obtenido con el ordenador en el contexto del problema
	PR15. Completar un contraste de hipótesis tomando una decisión.
	PR16. Manejo del software.
	PR17. Cálculo de los grados de libertad para realizar un test Chi-cuadrado.
	PR18. Manejo de tablas de distribuciones de probabilidad.
	PR19. Cálculo del estadístico de contraste.
Conceptos	C1. Variable simple / doble valor.
	C2. Distribución bidimensional
	C3. Variable cualitativa y cuantitativa.
	C4. Tablas de contingencia, tablas rxc.
	C5. Frecuencia absoluta y relativa doble
	C6. Frecuencia absoluta y relativa marginal; distribución marginal
	C7. Frecuencia relativa condicional; distribución condicional
	C8. Probabilidad simple.
	C9. Probabilidad conjunta y condicional.
	C10. Independencia /Asociación, dependencia funcional
	C11. Asociación inversa, directa.
	C12. Intensidad de la asociación.
	C13. Causalidad.
	C14. Diagrama de barras apilado.
	C15. Diagrama de barras adosado.
	C16. Gráfico tridimensional.
	C17. Frecuencias esperadas.
	C18. Estadístico Chi-cuadrado, grados de libertad, distribución.
	C19. Coeficiente Phi, Riesgo relativo, Razón de productos cruzados (medidas tablas 2x2).
	C20. Coeficiente C y V (medidas simétricas, tablas rxc).
	C21. Coeficiente Lambda (medidas asimétricas).
	C22. Significación.

	C23. <i>p</i> valor.	
	C24. Hipótesis alternativa y nula.	
Propiedades	PP1. Relaciones entre frecuencias absolutas y relativas.	
	PP2. Relaciones entre frecuencias dobles, condicionales y marginales.	
	PP3. Igualdad /diferencia de frecuencias relativas condicionales en caso de independencia/dependencia.	
	PP4. Igualdad/ diferencia de frecuencias relativas condicionales y marginales en caso de independencia/ dependencia.	
	PP5. Celdas que indican signo de asociación en la tabla 2x2.	
	PP6. Suma de frecuencias relativas igual a la unidad.	
	PP7. Frecuencias esperadas en caso de independencia.	
	PP9. Relación entre asociación y causalidad.	
	PP10. Diferencia entre dependencia aleatoria y funcional.	
	PP12. Relación entre valor <i>p</i> en el contraste Chi-cuadrado y asociación.	
	PP13. Diferencias entre el significado y aplicación de las medidas de asociación.	
	PP14. El diagrama de barras adosado se puede construir con frecuencias absolutas o porcentajes.	
	PP15. Relación y diferencia entre diagrama de barras adosado y apilado.	
	PP16. Propiedades de independencia.	
	PP17. Diferenciar valores de Chi-cuadrado estadísticamente significativos.	
	PP18. Condiciones de aplicación del estadístico Chi-cuadrado.	
	PP19. Relación entre el riesgo relativo por filas y columnas.	
	PP20. Rango de acción de las medidas de asociación.	
	PP21. Relación entre probabilidad y tabla de contingencia.	
	Argumentos	A1. Razonamientos informales.
		A2. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar.
A3. Deducción de propiedades sin excesiva formalización		

4.8. CONCLUSIONES SOBRE EL SIGNIFICADO INSTITUCIONAL PRETENDIDO

Una vez finalizado el análisis detallado de cada sesión teórica y práctica, estamos en condiciones de resumir en una Tabla (4.8.1) el significado institucional pretendido en el proceso de estudio diseñado. Un objetivo del trabajo era conseguir una adecuada idoneidad didáctica de dicho proceso de estudio, en sus diversas componentes, descritas por Godino, Contreras y Font (2006). Aunque algunos de estos componentes no se pueden valorar hasta el desarrollo del proceso si estamos en condiciones de realizar una valoración al menos potencial de algunos de ellos.

Idoneidad epistémica

En primer lugar trataremos de justificar la idoneidad epistémica del proceso diseñado, en base a que se logra la representatividad de los significados institucionales pretendidos, respecto al significado matemático identificado en el análisis incluido en el Capítulo 1 (Sección 1.4.1).

Tabla 4.8.1. Elementos de significado pretendido en el proceso de estudio

Tipo de Objetos	Objetos matemáticos	L1	L2	L3	L4	P1
Situaciones-problemas	P1. Resumir la información obtenida de un conjunto de observaciones bivariantes de modo que se visualicen posibles relaciones entre las variables.	X	X			X
	P2. Cálculo de probabilidades asociadas a valores específicos de una o las dos variables.	X				X
	P3. Contraste de homogeneidad.			X		X
	P4. Contraste de independencia.			X		X
	P5. Definir una medida de la intensidad de asociación entre las variables.				X	X
Lenguaje	R1. Diagrama de barras adosado.					X
	R2. Diagrama de barras apilado.	X				X
	R3. Gráfico tridimensional.	X				X
	R4. Lenguaje verbal.	X	X		X	X
	R5. Lenguaje simbólico.	X	X	X	X	X
	R6. Lenguaje gráfico.	X	X			
	R7. Tablas (de frecuencia y de contingencia).	X	X	X	X	X
Procedimientos	PR1. Lectura e interpretación de la tabla.	X	X	X	X	X
	PR2. Construcción de una tabla de contingencia.	X				X
	PR3. Cálculo de probabilidades simples, compuestas y condicionales a partir de la tabla.	X				X
	PR4. Cálculo de frecuencias condicionales y marginales (absolutas o relativas).	X	X			X
	PR5. Representación gráfica de datos a partir de una tabla de contingencia.	X				X
	PR6. Cálculo de las frecuencias esperadas.		X	X	X	X
	PR7. Comprobar las propiedades de independencia.		X		X	X
	PR8. Selección de medidas de la asociación.				X	X
	PR9. Cálculo e interpretación de medidas en tablas 2x2.				X	X
	PR10. Cálculo e interpretación de medidas en tablas rxc.				X	X
	PR11. Cálculo e interpretación de medidas basadas en la reducción del error.				X	X
	PR12. Formulación de hipótesis (contraste de independencia).			X		X
	PR13. Formulación de hipótesis (contraste de homogeneidad).			X		X
	PR14. Interpretación de resultados del test Chi-cuadrado obtenido con el ordenador en el contexto del problema.			X	X	X
	PR15. Completar un contraste de hipótesis tomando una decisión.			X		X
	PR16. Manejo del software.					X
	PR17. Cálculo de los grados de libertad para realizar un test Chi-cuadrado.			X		X
	PR18. Manejo de tablas de distribuciones de probabilidad.			X		X
	PR19. Cálculo del estadístico de contraste.			X	X	X
Conceptos	C1. Variable simple / doble valor.	X	X	X	X	X
	C2. Distribución bidimensional.	X	X	X	X	X
	C3. Variable cualitativa y cuantitativa.	X				X
	C4. Tablas de contingencia, tablas rxc.	X	X	X		X
	C5. Frecuencia absoluta y relativa doble.	X	X	X		X
	C6. Frecuencia absoluta y relativa marginal; distribución marginal.	X	X	X		X
	C7. Frecuencia relativa condicional; distribución condicional.	X	X	X		X
	C8. Probabilidad simple.	X				X
	C9. Probabilidad conjunta y condicional.	X				X
	C10. Independencia /Asociación, dependencia funcional.		X	X		X
	C11. Asociación inversa, directa.		X	X	X	X
	Objetos matemáticos	L1	L2	L3	L4	P1

	C12. Intensidad de la asociación.			X	X	X
	C13. Causalidad.		X			X
	C14. Diagrama de barras apilado.	X				X
	C15. Diagrama de barras adosado.	X				X
	C16. Gráfico tridimensional.	X				X
	C17. Frecuencias esperadas.		X	X	X	X
	C18. Estadístico Chi-cuadrado, grados de libertad, distribución.			X		X
	C19. Coeficiente Phi, Riesgo relativo, Razón de productos cruzados (medidas tablas 2x2).				X	X
	C20. Coeficiente C y V (medidas simétricas, tablas rxc).				X	X
	C21. Coeficiente Lambda (medidas asimétricas).				X	X
	C22. Significación.			X		X
	C23. p valor.			X		X
	C24. Contraste de hipótesis. Hipótesis alternativa y nula.			X		X
	C25. Población y muestra. Estadístico de contraste.			X		
	C26. Región crítica y de aceptación.			X		
	C27. Posibles errores en un contraste de hipótesis.			X		
Propiedades	PP1. Relaciones entre frecuencias absolutas y relativas.	X	X			X
	PP2. Relaciones entre frecuencias dobles, condicionales y marginales.	X	X			X
	PP3. Igualdad /diferencia de frecuencias relativas condicionales en caso de independencia/dependencia.		X			X
	PP4. Igualdad/ diferencia de frecuencias relativas condicionales y marginales en caso de independencia/ dependencia.		X			X
	PP5. Celdas que indican signo de asociación en la tabla 2x2.		X			X
	PP6. Suma de frecuencias relativas igual a la unidad.	X	X			X
	PP7. Frecuencias esperadas en caso de independencia.		X	X		X
	PP8. Las hipótesis de un contraste son complementarias.			X		
	PP9. Relación entre asociación y causalidad.		X			X
	PP10. Diferencia entre dependencia aleatoria y funcional.		X			X
	PP11. Relación entre regiones de aceptación y rechazo y decisión en el contraste.			X		
	PP12. Relación entre valor p en el contraste Chi-cuadrado y asociación.			X		X
	PP13. Diferencias entre el significado y aplicación de las medidas de asociación.				X	X
	PP14. El diagrama de barras adosado se puede construir con frecuencias absolutas o porcentajes.	X				X
	PP15. Relación y diferencia diagrama de barras adosadas y apiladas.	X				X
	PP16. Propiedades de independencia.		X			X
	PP17. Diferenciar valores de chi-cuadrado estadísticamente significativos.			X		X
	PP18. Posibles errores en un contraste.			X		
	PP18. Condiciones de aplicación del estadístico Chi-cuadrado.			X		X
	PP19. Relación entre el riesgo relativo por filas y columnas.				X	X
	PP20. Rango de acción de las medidas de asociación.				X	X
	PP21. Relación entre probabilidad y tabla de contingencia.	X			X	X
PP22. Suma de frecuencias absolutas dobles igual al total de la muestra.	X					
Argum entos	A1. Razonamientos informales.	X	X	X	X	X
	A2. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar; comparación de frecuencias.	X	X	X	X	X
	A3. Deducción de propiedades sin excesiva formalización.					X

- El proceso de estudio considera cuatro de los cinco campos de problemas relacionados con las tablas de contingencia. Únicamente se ha dejado sin incluir el ajuste de modelos matemáticos en tablas de contingencia (modelos log-lineales) que requeriría un conocimiento estadístico mucho más avanzado que el disponible por los estudiantes.
- En cuanto al *lenguaje*, a lo largo de las lecciones se incluirán distintos tipos de lenguajes verbal, simbólico, gráficos y tablas. Se comienza con el uso del lenguaje verbal, introduciendo de forma gradual el resto.
- Todos los *procedimientos* identificados en el estudio del significado de referencia se presentan en el proceso de estudio: lectura, construcción e interpretación de tablas, probabilidades, representaciones gráficas, cálculo de los distintos tipos de frecuencias, contraste de independencia, contraste de homogeneidad, cálculo e interpretación de medidas de asociación, manejo de software, manejo de tablas de distribución Chi-cuadrado y comprobación de propiedades de independencia.
- Los *conceptos* son variados, y se incluyen todos los esenciales, ya que por ejemplo la primera lección se explican todos los elementos que conforman las tablas de contingencia. Además, el hecho de que los estudiantes tenga una cierta base en estadística (incluido los contrastes de hipótesis), permitirá el uso de un número más amplio de conceptos.
- Las *propiedades* al igual que los procedimientos, son los identificados en el estudio del significado, destacando el hecho que se puedan introducir todas las propiedades sin grandes dificultades.
- Los *argumentos* que aparecen en la enseñanza son dos: razonamientos informales, y comprobación de ejemplos y contraejemplos, sin pretensión de generalizar. No se pudieron hacer demostraciones formales, por lo mencionado anteriormente de que no se les exige saber formulas en la asignatura, y mucho menos demostraciones teóricas.

En resumen, esta enumeración muestra una razonable *idoneidad epistémica*, en cuanto que, aunque el significado institucional pretendido es más limitado que el de referencia, contempla los elementos de significado más importantes de las tablas de contingencia.

Puesto que la enseñanza es adaptada al tiempo disponible y conocimientos previos de los estudiantes se prevé una buena *idoneidad cognitiva*. Además, la propuesta de enseñanza propone la introducción en forma gradual del manejo de tablas de contingencia, de manera que el estudiante llegue al final de la enseñanza sabiendo un manejo bastante amplio de este instrumento, poniendo al estudiante en situaciones de reflexión desde el punto de vista teórico (mediante clases tradicionales), práctico (donde realicen operaciones) e incluso de interpretación (donde el software les dice los resultados y ellos interpretan). Esto implica el manejo de software del procesador de texto Word y la hoja de cálculo Excel, además del uso de Internet. También habría que mencionar que los estudiantes no tienen que memorizar formulas, ya que se les permiten en el examen un formulario con todas las que han visto.

Se hace un uso de diversos recursos didácticos (a través de las configuraciones epistémicas), que junto a los momentos previstos de evaluación a lo largo del proceso de estudio contribuye a su *idoneidad didáctica*. Finalmente, puesto que los campos de problemas se contextualizan en aplicaciones en el área de psicología como agentes motivadores de su rol profesional la propuesta tiene una *idoneidad afectiva*. Por supuesto estos tipos de idoneidad son previstos a priori y habrá que contrastarlo con la implementación efectiva y aprendizaje de los estudiantes para analizar la idoneidad final del proceso de estudio. Esta tarea se lleva a cabo en los capítulos 5 y 6.

CAPÍTULO 5.

EVALUACIÓN DEL APRENDIZAJE MEDIANTE CUESTIONARIOS

- 5.1. Introducción
- 5.2. Objetivos y características de los cuestionarios
 - 5.2.1. Objetivos
 - 5.2.2. Proceso de construcción
- 5.3. Análisis a priori del significado evaluado en el cuestionario
 - 5.3.1. Análisis del cuestionario 1
 - 5.3.2. Análisis del cuestionario 2
 - 5.3.3. Análisis de la prueba final
- 5.4. Síntesis de las evaluaciones mediante cuestionario
- 5.5. Resultados en el cuestionario
 - 5.5.1. Resultados del primer cuestionario
 - 5.5.2. Resultados en el segundo cuestionario
 - 5.5.3. Síntesis de resultados en los cuestionarios
- 5.6. Resultados en la prueba final
- 5.7. Conclusiones

5.1. INTRODUCCIÓN

Durante el desarrollo de la experiencia de enseñanza se propuso a los estudiantes de la muestra varias pruebas para evaluar los conocimientos adquiridos, utilizando cuestionarios y problema abiertos. Por un lado, se dedicaron dos sesiones de clase (de 1 hora de duración) a responder dos cuestionarios compuestos en total por 24 ítems, cada uno de los cuáles contenía 3 o 4 subítems de verdadero/falso (6 de cada tema). Asimismo, como parte del examen final de la asignatura, se propusieron 4 ítems de opción múltiple (con cuatro opciones, de las cuales solo una es respuesta).

Por otro lado, también en sesiones de clase se propusieron varios problemas abiertos, algunos de los cuales se resolvieron con ayuda del programa en Excel desarrollado para facilitar el trabajo con las tablas de contingencia.

En base a estos instrumentos, analizamos en los Capítulos 5 y 6 la comprensión lograda por los estudiantes de los diferentes elementos de significados relacionados con la tabla de contingencia, incluidos en el proceso de estudio descrito en el Capítulo 4. Para este trabajo, se entiende la evaluación como un estudio de la correspondencia entre el *significado institucional* presentado en la enseñanza y el *significado personal*

efectivamente alcanzado por los estudiantes (Godino, 2002b). Es decir, evaluamos el uso correcto del lenguaje, comprensión de conceptos y propiedades, aplicación de procedimientos y capacidad de argumentación. El resultado de la evaluación servirá también para valorar la idoneidad cognitiva del proceso de estudio seguido por los estudiantes (Godino, Contreras y Font, 2006).

En este capítulo se realiza el análisis de los resultados recogidos mediante los ítems de opción múltiple en los diferentes cuestionarios y en el Capítulo 6, los obtenidos de los problemas abiertos propuestos. En lo que sigue se analizarán los ítems que componen los cuestionarios y presentarán los resultados.

5.2. OBJETIVOS Y CARACTERÍSTICAS DE LOS CUESTIONARIOS

5.2.1. OBJETIVOS

El objetivo principal al construir los cuestionarios fue disponer de instrumentos que, en un corto espacio de tiempo, permitiera recoger datos que sirviesen para aproximarnos a la comprensión mostrada en el grupo de estudiantes, en relación con los principales elementos de significado incluidos en el proceso de estudio. De este objetivo principal se deducen otros específicos:

- El *primer objetivo específico* fue estimar la proporción de estudiantes en la muestra que comprenden y aplican los objetos matemáticos incluidos en la enseñanza.
- Asimismo, como *segundo objetivo*, se quiso identificar las dificultades presentadas respecto a dichos objetos, mostrando las características en la comprensión en el grupo de estudiantes, tanto en sus tendencias, como su variabilidad.

Con estos objetivos queremos conseguir el objetivo O5 de esta investigación expuesto y comentado en la sección 1.7 del Capítulo 1.

Todo ello permitirá caracterizar el significado personal de los estudiantes sobre las tablas de contingencia y compararlo con el mostrado en el estudio inicial de evaluación (Capítulo 3), destacando la posible superación de los conflictos semióticos descritos en el citado capítulo, así como la adquisición de nuevos elementos de significado formales del concepto de asociación. Posteriormente, en el Capítulo 6 se completa esta evaluación con el análisis de los problemas abiertos, que nos permitirá acceder a un nivel más completo de los razonamientos de los estudiantes e identificar sus conflictos semióticos.

5.2.2. PROCESO DE CONSTRUCCIÓN

En el proceso de elaboración de los cuestionarios se siguieron una serie de recomendaciones psicométricas para asegurar la calidad de los instrumentos (Osterlind, 1989; Thorndike, 1989; Martínez, 1995; Martínez, Hernández, Hernández, 2006):

- En primer lugar, se delimitó el contenido a evaluar con estos instrumentos, a partir del análisis del estudio del significado de referencia llevado a cabo en la Sección 1.5 y del análisis del proceso de estudio diseñado, que se describe en el Capítulo 4.
- Se especificó el formato de los ítems, decidiendo incluir ítems de verdadero/ falso, que permiten en poco tiempo acceder a la evaluación de muchos objetos matemáticos simultáneamente, aunque con menor profundidad. Para conseguir evaluar de manera más profunda los argumentos y procedimientos que usan los estudiantes, así como, en general, su comprensión, se propuso a los estudiantes algunos problemas abiertos que se analizan en el Capítulo 6 y que permiten valorar habilidades de pensamiento de alto nivel (Gal, 2002).
- El número de ítems (6 ítems para cada tema, en total 24) y problemas (problemas que abarcan todos los temas) se fijó teniendo en cuenta el tiempo disponible, el tipo de evaluaciones a las que estaban acostumbrados los estudiantes y que se deseaba cubrir el máximo de objetos matemáticos.
- Se procedió a la elaboración de una colección de ítems iniciales. A partir de ellos, mediante una serie de revisiones por parte del doctorando y los directores, así como de dos profesores del área de metodología de las ciencias del comportamiento, y analizando los contenidos evaluados, se seleccionaron, posteriormente, los que habrían de constituir los cuestionarios.

Algunos de los ítems y problemas son de elaboración propia y otros han sido tomados o adaptados de diferentes libros de texto o investigaciones previas. Una vez que se llegó a un conjunto de ítems que cubría el contenido pretendido, se procedió a homogeneizar la redacción y a cambiar el contexto cuando éste no fuese familiar al estudiante. Se hicieron pruebas de comprensión del enunciado, con algunos estudiantes de la misma especialidad, modificando la redacción en los casos que fue necesario.

En la redacción de los enunciados se tuvieron en cuenta los siguientes aspectos: Evitar detalles innecesarios, relevancia de las preguntas formuladas para el estudio, nivel de lectura adecuado, brevedad, evitar las cuestiones negativas, evitar cuestiones sesgadas o interdependientes, claridad y falta de ambigüedad, que la respuesta sea

razonable para el sujeto y pueda darla, evitar hipótesis implícitas, nivel apropiado de abstracción, asegurar que las preguntas tienen el mismo significado para todos los sujetos.

5.3. ANÁLISIS A PRIORI DEL SIGNIFICADO EVALUADO EN EL CUESTIONARIO

Los cuestionarios presentan 24 ítems, que se describen a continuación, analizando su contenido y los posibles errores que, a priori, es previsible encontrar en las respuestas de nuestros estudiantes. Posteriormente se analizan los ítems incluidos en la prueba final de la asignatura.

El conjunto de ítems se dividió en dos cuestionarios (el primero con 12 ítems de los temas 1 y 2 y el segundo con 12 ítems de los temas 3 y 4) con objeto de pasarlos en dos sesiones diferentes, de modo que se evitase el cansancio en los estudiantes. Se incluyeron seis ítems para cada uno de los temas planificados en el proceso de estudio (Anexo 3). Denotaremos cada uno de ellos como X_n , donde X indica el tema al que corresponde el ítem (A indica tema 1, B indica tema 2, C indica tema 3 y D indica tema 4) y n el número de orden de la pregunta.

Los ítems se componen de un enunciado inicial y varias opciones de verdadero o falso, cada una de las cuáles puede ser correcta o no. Se marcan en negrita las respuestas correctas. Se dio instrucciones a los estudiantes sobre la forma de responder al cuestionario, indicándoseles que cada uno de los sub-ítems podía ser verdadero o falso y pidiéndoles que lo completasen con cuidado e interés.

A continuación se analiza cada cuestionario.

5.3.1. ANÁLISIS DEL CUESTIONARIO 1

Análisis del ítem 1 del cuestionario 1

- A1. La principal diferencia entre el diagrama de barras apilado y el adosado es:
- a. El diagrama de barras apilado sólo se puede construir con frecuencias absolutas, mientras que el adosado se puede construir con frecuencias absolutas o porcentajes
 - b. En el diagrama de barras apilado las frecuencias condicionales de cada valor de X para un mismo valor de Y se representan en una misma barra y en el adosado se representan una al lado de la otra.**
 - c. En el diagrama de barras apilado las frecuencias marginales de cada valor de X se representan en una misma barra y en el adosado se representan una al lado de la otra

Este ítem evalúa la comprensión que tienen los estudiantes de las diferencias entre dos de las representaciones posibles de los datos de una tabla de contingencia: El diagrama de barras adosado y apilado. La respuesta correcta es la b), pues el objeto de los dos diagramas es representar las distribuciones condicionales de una de las variables (por ejemplo la X) respecto a cada uno de los valores de la otra (en este caso la Y). Para responder la pregunta, el estudiante, además de los convenios de construcción de gráficos, necesita conocer el concepto de frecuencia condicional y marginal.

- La respuesta a) es incorrecta, pues ambos diagramas pueden construirse con frecuencias absolutas, relativas o porcentajes.
- Asimismo lo es la respuesta c), porque lo que se representa son frecuencias condicionales y no marginales.

Análisis del ítem 2 del cuestionario 1

A2. En una tabla de contingencia, la suma de las frecuencias relativas marginales por filas es:

- a. **Igual a la suma de frecuencias relativas marginales por columnas**
- b. Igual al total de la muestra
- c. **Igual a la suma de las frecuencias absolutas marginales dividida por el total de la muestra**

Este ítem es uno de los dedicados a evaluar la comprensión que el estudiante adquiere de la relación entre los diversos tipos de frecuencia en una tabla de contingencia, Puesto que los estudiantes confunden a veces las frecuencias condicionales con su transpuesta o bien con las frecuencias simples y compuestas (Falk, 1986; Ojeda, 1995; Contreras, 2011), es posible que esta confusión se extienda a los diferentes tipos de frecuencia: En este caso, las frecuencias relativas marginales por filas y columnas; la relación de las frecuencias relativas marginales, con el total de la muestra; ó la existente entre las frecuencias relativas marginales, y frecuencias absolutas marginales (Nortes, 1993).

La respuesta a) es correcta, pues una de las propiedades que tienen que cumplir las distribuciones relativas, incluidas las marginales, es que la suma de sus frecuencias sumen 1, por tanto la suma de las frecuencias relativas marginales por filas es igual a la suma de las frecuencias relativas marginales por columnas; la c) es también correcta, ya que las frecuencias relativas marginales se calculan dividiendo las frecuencias absolutas marginales por el tamaño de la muestra.

- La respuesta b) se debe rechazar, pues el tamaño de la muestra se obtiene al sumar la distribución absoluta marginal, y no las distribuciones relativas marginales, que

suman el valor 1. El estudiante tiene que distinguir entre distribuciones absolutas y marginales, para identificar que esta opción es falsa.

Análisis del ítem 3 del cuestionario 1

A3. Las frecuencias dobles absolutas y relativas están relacionadas entre sí mediante:

- a. Las frecuencias condicionales.
- b. Las frecuencias marginales.
- c. **El tamaño de la muestra.**

Se estudia la comprensión de la relación de las frecuencias dobles absolutas y relativas en una tabla de contingencia, lo que básicamente requiere recordar su formulación y cálculo (Nortes, 1993). La respuesta correcta es la c), ya que las frecuencias dobles relativas se obtienen al dividir las frecuencias dobles absolutas por el tamaño de la muestra.

- La respuesta a) es incorrecta, puesto que, aunque las frecuencias condicionales se obtienen dividiendo las frecuencias dobles absolutas ó relativas por la frecuencia marginal absoluta ó relativa, no se pueden utilizar para establecer una relación entre las frecuencias dobles absolutas y relativas.
- Otro error sería elegir la respuesta b), ya que las frecuencias marginales se obtienen sumando las frecuencias dobles absolutas ó relativas, pero no relacionan las frecuencias dobles absolutas y relativas.

Análisis del ítem 4 del cuestionario 1

A4. Para calcular las frecuencias relativas condicionadas es necesario:

- a. **Las frecuencias absolutas dobles de la fila o columna por la que se condiciona.**
- b. Las frecuencias relativas de la fila o columna por la que se condiciona y el total de la muestra.
- c. **Las frecuencias absolutas o relativas de la fila o columna por la que se condiciona.**

Este ítem considera la relación de las frecuencias relativas condicionales con algunos componentes de la tabla de contingencia (Amón, 1993) y se ha elegido porque las frecuencias condicionales, al igual que las probabilidades condicionales producen bastante dificultad para los estudiantes (Contreras, 2011). Una respuesta correcta es la a), pues la frecuencia condicionada se obtendría dividiendo las frecuencias absolutas o relativas de la fila o columna por la que se condiciona, por la frecuencia marginal de esta fila o columna. Por el mismo motivo, la respuesta c), también es correcta.

- La respuesta b) es errónea, pues no se requiere el uso del total de la muestra. El

estudiante tiene que comprender el cálculo de las frecuencias condicionadas y la diferencia entre los distintos tipos de frecuencias para responder el ítem.

Análisis del ítem 5 del cuestionario 1

A5. Las tablas rxc se pueden representar gráficamente en los gráficos:

- a. Únicamente en el diagrama de barras apilado y el adosado.
- b. Únicamente en el gráfico tridimensional y el diagrama de barras adosado.
- c. **En los mismos gráficos que una tabla 2x2.**

Este ítem evalúa la comprensión de la relación del tamaño de la tabla de contingencia, con los distintos tipos de gráficos que se pueden utilizar. Una tabla rxc tiene la propiedad de poder subdividirse en subtablas 2x2 (Ato y López, 1996), así que podría considerarse la representación gráfica de una tabla rxc de un nivel superior que la representación de una tabla 2x2. Como consecuencia, la respuesta correcta es la c), ya que no existen unos gráficos específicos para las tablas 2x2. En cualquier tabla se pueden utilizar el diagrama de barras apilado, el diagrama de barras adosado y el gráfico tridimensional. Para responder correctamente la pregunta, el estudiante, además de la construcción de gráficos, necesita conocer cuando se pueden aplicar estos gráficos.

- La respuesta a) es incorrecta, ya que las tablas rxc además de poderse representar en el diagrama de barras apilado y el adosado, pueden representarse en el gráfico tridimensional,
- También lo es la respuesta b), ya que las tablas rxc además de poderse representar en el diagrama de barras adosado y el gráfico tridimensional, pueden representarse en el diagrama de barras apilado.

Análisis del ítem 6 del cuestionario 1.

A6. El objetivo de construir una tabla de contingencia es:

- a. Resumir los datos de dos variables cuantitativas
- b. **Resumir los datos de dos variables cualitativas**
- c. **Resumir los datos de dos variables cualitativas, o bien numéricas con pocos valores diferentes**

Este ítem evalúa los conocimientos del estudiante sobre los convenios de construcción de las tablas de contingencia, y el objetivo de estas tablas. Un concepto importante que el estudiante tiene que tener claro son los tipos de variables (cuantitativo y cualitativo) y cuáles son adecuados a estas tablas. La respuesta correcta es la b), ya que el objetivo es resumir los datos, de manera que se pueda trabajar con ellos, y este

tipo de tablas se emplean para variables cualitativas. La respuesta c) es también correcta, aunque el uso de las tablas es más generalizado en variables cualitativas.

- La respuesta a) es errónea, pues en las tablas de contingencia se representan preferentemente variables cualitativas, y no cuantitativas.

Análisis del ítem 7 del cuestionario 1

B1. 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.**

Los ítems siguientes evalúan el conocimiento que el estudiante adquiere de propiedades sencillas que permiten evaluar si dos variables de una tabla de contingencia son independientes (Batanero y Díaz, 2008). En este caso, una respuesta correcta es la a), pues en caso de independencia, todas las distribuciones condicionales por columna coinciden con la distribución marginal de la variable X o, lo que es lo mismo, la distribución de X no cambia cuando se condiciona por un valor de Y .

La opción b) es también adecuada, ya que en caso de independencia, todas las distribuciones condicionales por fila coinciden con la distribución marginal de la variable Y o lo que es lo mismo, la distribución de Y no cambia cuando se condiciona por un valor de X . La c) es igualmente correcta, ya que, como hemos dicho, la independencia implica esta relación de igualdad entre las frecuencias condicionales y las frecuencias marginales. Para responder la pregunta, el estudiante tiene que conocer estas propiedades de la independencia.

Análisis del ítem 8 del cuestionario 1

B2. 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

Mediante este ítem se desea evaluar el conocimiento del estudiante del procedimiento de cálculo de las frecuencias esperadas en una tabla de contingencia. La respuesta correcta es la c), pues para calcular estas frecuencias, hay que multiplicar las frecuencias marginales entre sí, y dividir este producto por el total de la muestra, es decir:

$e_{ij} = \frac{f_{i.} \times f_{.j}}{n}$. Para responder la pregunta, el estudiante tiene que diferenciar los

distintos tipos de frecuencia en las tablas y recordar el proceso de cálculo de las frecuencias esperadas.

- La respuesta a) es incorrecta, ya que confunde las frecuencias absolutas dobles con las frecuencias absolutas marginales, en el cálculo de las frecuencias esperadas.
- La respuesta b) se debe rechazar, pues cambia las frecuencias relativas dobles por las frecuencias absolutas marginales, en el proceso de cálculo de las frecuencias esperadas.
- Otro error se recoge en la respuesta d), consistente en confundir las frecuencias relativas marginales con las frecuencias absolutas marginales, además de que falta el total de la muestra, en el cálculo.

Análisis del ítem 9 del cuestionario 1

B3. 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 $h_{i,j} = h_{i.} \cdot h_{.j}$

Este ítem evalúa el conocimiento del estudiante de la relación entre frecuencias relativas y marginales en caso de independencia en una tabla de contingencia (Batanero y Díaz, 2008). El estudiante tiene que distinguir entre propiedades de independencia y asociación, para identificar las opciones verdaderas y falsas. La respuesta correcta es la b), pues en caso de asociación podría darse la igualdad en un caso pero no en todas las celdas.

- 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.

Análisis del ítem 10 del cuestionario 1.

B4. 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

Mediante este ítem 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. Las celdas (presencia-presencia y ausencia-ausencia) informarían que la asociación en la tabla es directa, según Inhelder y Piaget (1955). 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.

- La respuesta a) es incorrecta, porque cuando se da un solo carácter y el otro no, la celda informa de una dependencia inversa, además cuando hay ausencia-ausencia la celda informa de una dependencia directa.
- Un error posible se recoge en c), ya que cuando no se da ninguno de los caracteres la celda informa de una dependencia inversa. Este distractor se elige porque en investigaciones (Smedlund, 1965; Beyth y Maron, 1982; Shaklee y Mins, 1982; Yates y Curley, 1986; Estepa, 1993) se ha comprobado que muchos individuos basan sus juicios de asociación directa únicamente en la celda *a*.

Análisis del ítem 11 del cuestionario 1

B5. Indica cuál de las siguientes frases es cierta:

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

Este ítem evalúa la comprensión del estudiante de las posibles explicaciones de la asociación, descritas por Barbancho (1973) y también trata de detectar la concepción

causal de la asociación propuesta por Estepa (1993), por lo que, para resolverlo el alumno ha de comprender que aunque la causalidad implica correlación, la correlación no implica causalidad. La respuesta adecuada es la c), pues si A es causa de B , entonces habrá asociación, ya que lo que pase en A influye en B de forma positiva o negativa.

- 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”, y estamos olvidando el caso de un tercer factor, u otras posibles explicaciones, como la concordancia o interdependencia (Barbancho, 1973).

Análisis del ítem 12 del cuestionario 1

B6. La diferencia entre la dependencia funcional y la dependencia aleatoria consiste en:

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

Este ítem trata de analizar la comprensión de la diferencia entre dependencia funcional y dependencia aleatoria entre dos variables (Nortes, 1993). La respuesta correcta es la b), pues la dependencia aleatoria no implica un único valor de la variable dependiente al variar la independiente, lo que produce que al variar X pueda variar Y , puesto que sólo conocemos una distribución de valores.

- La respuesta a) es incorrecta, pues es la dependencia funcional cuando ocurre que X le corresponde sólo un valor de la variable Y , para todos los casos, mientras en la aleatoria corresponde una distribución de valores.
- La respuesta c) es asimismo incorrecta, puesto que la dependencia funcional no es siempre directa, puede ocurrir que sea inversa.

5.3.2. ANÁLISIS DEL CUESTIONARIO 2

Análisis del ítem 1 del cuestionario 2

C1. Para realizar un test Chi-cuadrado en una tabla de contingencia con m filas y n columnas, los grados de libertad del estadístico Chi-cuadrado son:

a. $m \times n$.	b. $m + n$.
c. $(m-1) \times (n-1)$.	d. $(m-1) + (n-1)$.

Este ítem evalúa el conocimiento del cálculo de los grados de libertad en el estadístico Chi-cuadrado, calculado a partir de una tabla de contingencia. La respuesta adecuada es la c), pues los grados de libertad de un estadístico calculado sobre un conjunto de datos se refieren al número de cantidades independientes que se necesitan en su cálculo, menos el número de restricciones que ligan a las observaciones y el estadístico (Martín y Luna, 1994). Se calcula, en primer lugar el número de sumandos, es decir (número de filas) x (número de columnas) en la tabla. A esta cantidad se debe restar el número de restricciones impuestas a las frecuencias observadas, por tanto los grados de libertad son $(m-1) \times (n-1)$.

- La respuesta a) corresponde a un error, pues no se resta el número de restricciones que ligan a las observaciones y el estadístico. El error de no restarlas apareció en el cálculo de grados de libertad ligados a algunas distribuciones de probabilidad en el trabajo de Alvarado (2007).
- La respuesta b) tampoco es correcta, pues no se tiene en cuenta que se resta el número de restricciones que ligan a las observaciones y el estadístico, además se está confundiendo el producto con la suma, es decir no se calcula bien el número de sumandos.
- La respuesta d) es incorrecta, pues se está confundiendo el producto con la suma.

Análisis del ítem 2 del cuestionario 2

C2. Los grados de libertad sirven para:

a. Contar el número de filas y columnas.

b. Calcular el parámetro de la distribución del estadístico.

c. Encontrar en las tablas de la distribución la probabilidad de un valor de Chi-cuadrado, y saber si el valor de Chi cuadrado es estadísticamente significativo.

Este ítem evalúa la comprensión del significado de los grados de libertad en un estadístico Chi-cuadrado. Las respuestas correctas son dos: en primer lugar la b), pues los grados de libertad son el parámetro de la distribución del estadístico (Martín y Luna, 1994). También es correcta la c), ya que los grados de libertad se utilizan para buscar en las tablas de la distribución el valor crítico, que nos informa de si el valor obtenido de Chi-cuadrado es estadísticamente significativo. Para responder la pregunta, el estudiante tiene que conocer el significado de los grados de libertad y su uso en las tablas de la distribución.

- La respuesta a) es incorrecta, pues esta respuesta implica que el estudiante confunde el cálculo de los grados de libertad, con su utilidad.

Análisis del ítem 3 del cuestionario 2

C3. Un investigador usa un test Chi-cuadrado para determinar si hay asociación entre dos variables en una tabla de contingencia. Calcula el valor de Chi-cuadrado y obtiene el valor p correspondiente. ¿Cuál de los siguientes valores p indicaría de forma más concluyente que hay asociación entre las variables?.

- | | |
|----------------|--------|
| a. .002 | b. .01 |
| c. .05 | d. .10 |

Este ítem evalúa la interpretación del valor p o probabilidad de obtener un valor igual o más extremo del estadístico en una muestra aleatoria del mismo tamaño (Lipschutz y Schiller, 1999). La respuesta esperada es la a), pues un valor $p=0$, indicaría asociación perfecta, y valor $p=1$, nos llevaría a concluir independencia total; esto quiere decir que el valor más pequeño indicaría de forma más concluyente que hay asociación entre las variables, siendo este valor en nuestro caso de 0.002.

- Las respuestas b), c) y d) son incorrectas, pues no son los valores menores de p . El estudiante tiene que saber, que un valor de p alto es indicativo independencia, y un valor de p muy bajo apoyaría la asociación.

Análisis del ítem 4 del cuestionario 2

C4. Un investigador quiere estudiar si existe relación entre dos variables. Calcula el estadístico Chi-cuadrado y obtiene un valor p estadísticamente significativo ($p < 0,05$). La decisión que debe tomar es:

- Rechazar la hipótesis de independencia entre las dos variables**
- Rechazar la hipótesis de asociación entre las dos variables
- Rechazar la hipótesis de asociación directa entre las dos variables
- Con estos datos, no podemos saber la decisión que se debe tomar

Mediante este ítem se desea conocer la interpretación que dan los alumnos al valor p en una situación real, descrita en el enunciado, así como su relación con las hipótesis (Lipschutz y Schiller, 1999). La respuesta que se espera es la a) Pues, por convenio, como $p < 0,05$ habría que rechazar la hipótesis de independencia. Para responder la pregunta, el estudiante tiene que conocer el valor p , el significado del nivel de significación 0,05, además de su relación con la decisión de aceptación o rechazo de la hipótesis.

- La respuesta b) es incorrecta, pues un valor p menor que el nivel de significación sería estadísticamente significativo y en este caso habría que rechazar la hipótesis de independencia, aceptando la existencia de asociación.
- También lo es la respuesta c), pues, aunque el valor p lleva al rechazo de la independencia, no es posible con esta información diferenciar entre asociación directa o inversa.

- Asimismo es errónea la respuesta d), pues tenemos información suficiente, ya que por convenio sabemos que se ha de rechazar la hipótesis de independencia para un valor p inferior a 0,05.

Análisis del ítem 5 del cuestionario 2

C5. Un investigador quiere estudiar la relación entre dos variables estadísticas. Usando un programa de ordenador obtiene los siguientes valores en un contraste Chi-cuadrado

Pruebas de Chi-cuadrado			
	Valor	gl	Significación (valor p)
Chi-cuadrado de Pearson	157,331(a)	4	,005
N de casos válidos	398		

(a) 0 casillas (0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 14,75.

- El contraste es estadísticamente significativo, pero el investigador no puede aplicar el contraste Chi-cuadrado porque ninguna celda tiene frecuencia esperada menor que 5
- El contraste es estadísticamente significativo, y el investigador puede aplicar el contraste Chi-cuadrado porque se cumplen las condiciones**
- El contraste es estadísticamente significativo, pero el investigador no puede aplicar el contraste Chi-cuadrado porque la frecuencia mínima esperada es menor que 20.

Este ítem evalúa la comprensión de las condiciones de aplicabilidad del test Chi-cuadrado (Batanero y Díaz, 2008). La respuesta correcta es la b), pues se cumplen las condiciones de que, como mucho el 20% de las frecuencias esperadas sean menores que el valor 5 (en este caso no hay ningún caso con un valor menor que 5), y que ninguna frecuencia esperada sea inferior a 1 (La frecuencia mínima esperada es 14,77).

- Un error sería elegir a), pues no hay frecuencias esperadas ser menores que el valor 5.
- Asimismo es incorrecto elegir c), pues la frecuencia mínima esperada en este caso es 14,77.

Análisis del ítem 6 del cuestionario 2

C6. El investigador anterior:

- Debe concluir que no hay asociación entre las dos variables en su estudio
- Debe concluir que hay asociación entre las dos variables en su estudio, y la probabilidad de que las variables en realidad sean independientes es menor que 0,005
- Debe concluir que hay asociación entre las dos variables en su estudio; la probabilidad de obtener el valor Chi= 157, 331 si las variables son en realidad independientes es menor que 0,005**

Mediante este ítem se quiere estudiar el significado que los estudiantes asignan al valor p en una situación real, y si comprenden su relación con la decisión sobre el rechazo o aceptación de la hipótesis (Batanero y Díaz, 2008). Para responderlo debe

conocer la lógica del contraste de hipótesis y del contraste de independencia, en particular, así como comprender los conceptos de valor p , nivel de significación, región crítica y de aceptación y error tipo I. Se espera que el alumno elija la opción c), pues el valor p obtenido es estadísticamente significativo (menor que el nivel de significación), por lo que debe rechazar la hipótesis de independencia y aceptar la asociación entre las variables. Además este ítem incluye la definición de la probabilidad de error tipo I o probabilidad del error si se rechaza la hipótesis nula cuando en verdad es cierta.

- La respuesta a) es incorrecta, pues al ser el valor p bajo, se debe rechazar la hipótesis nula
- También lo es la respuesta b), pues aunque la primera parte da una regla de decisión adecuada, sobre el rechazo de la independencia, la definición que incluye de probabilidad de error es incorrecta, ya de hecho, no se puede calcular la probabilidad de una hipótesis en un contraste, un error descrito por Vallecillos (1994).

Análisis del ítem 7 del cuestionario 2

- D1. Si el coeficiente Phi tiene un valor 0,3, esto indica
- a. Relación inversa y moderada entre las variables
 - b. Relación directa y moderada entre las variables**
 - c. Relación directa y alta entre las variables
 - d. Relación inversa y alta entre las variables

Este ítem evalúa la interpretación que da el alumno a un valor del coeficiente de asociación Phi. La respuesta esperada es la b), pues este coeficiente toma valores entre 1 y -1, siendo la relación directa entre 0 y 1 (como es nuestro caso de 0,3), y siendo moderada en este caso por estar por debajo de 0,5 (Ato y López, 1996).

- Las respuestas a) y d) son incorrectas, pues al ser el signo positivo, la relación sería directa y no inversa. El estudiante tiene que distinguir entre la interpretación en el intervalo $[-1,0)$ y la interpretación en el intervalo $(0,1]$, para identificar que esta opción es falsa.
- Otro error se recoge en c), pues la relación es moderada en este caso por estar por debajo de 0,5. El estudiante ha de diferenciar los valores indicativos de asociación moderada y alta.

Análisis del ítem 8 del cuestionario 2

D2. Las medidas de la asociación para tablas 2x2 son:

- a. Coeficiente Phi de Pearson, riesgo relativo y coeficiente de contingencia de Pearson.
- b. Coeficiente Phi de Pearson, riesgo relativo y razón de productos cruzados.**
- c. Riesgo relativo, razón de productos cruzados y V de Cramer.

Con este ítem se desea analizar la competencia del estudiante para elegir unas medidas de asociación u otras, dependiendo del tipo de tabla de contingencia (Pardo y San Martín, 1998). La respuesta adecuada es la b), pues el coeficiente Phi de Pearson, riesgo relativo y razón de productos cruzados son medidas de asociación en tablas 2x2.

- La respuesta a) es incorrecta, pues el coeficiente de contingencia de Pearson es una medidas de asociación calculable sólo en tablas rxc.
- Otra posible respuesta incorrecta es c), pues el valor V de Cramer es una medidas de asociación adecuada únicamente para tablas rxc.

Análisis del ítem 9 del cuestionario 2

D3. El riesgo relativo por filas y columnas:

- a. Difieren y el valor por filas siempre es mayor o igual que el valor por columnas.
- b. Son diferentes su forma de cálculo, pero coinciden.
- c. Difieren y no hay una regla para saber cual es mayor.

Este ítem evalúa la interpretación del riesgo relativo por filas y por columnas. Para responder la pregunta, el estudiante tiene que conocer su significado como coeficiente que indica la disminución en el error relativo al predecir una variable en función de otra, y comprender que esta medida de asociación es asimétrica (Ato y López, 1996). En este ítem todas las opciones son falsas, como analizamos a continuación:

- La respuesta a) lo es porque el riesgo relativo por columnas indica cuanto más probable es la presencia de A con B que entre aquellos que no poseen B ; y el riesgo relativo por filas indica cuanto más probable es la presencia de B con A que entre aquellos que no poseen A . Estos dos valores puede que coincidan (sin ser lo habitual), pero también podría ser el primero sea mayor que el segundo ó al contrario.
- Asimismo lo son es respuesta b), pues su forma de cálculo es diferente, pero estos dos valores puede que coincidan (sin ser lo habitual) y la respuesta c) también es incorrecta, pues estos dos valores puede que coincidan.

Análisis del ítem 10 del cuestionario 2

D4. La razón de productos cruzados varía de:

- $0 < RC < 1$.
- $-1 < RC < 1$.
- Puede ser mayor o menor que 1**

Este ítem evalúa la comprensión de la razón de productos cruzados y el conocimiento de los distintos valores que puede tomar, que no están restringidos al intervalo (-1,1) (Ato y López, 1996). La respuesta correcta es la c), ya podía ocurrir que $RC = 1$, si hay la misma razón de casos que aparece A y \bar{A} , cuando está B presente, que cuando no lo está; también es posible que $RC < 1$, cuando la razón entre los casos que aparecen A y \bar{A} es menor cuando está presente B , y que $RC > 1$, si que la razón entre los casos que aparecen A y \bar{A} es mayor cuando está presente B .

- Las respuestas a) y b) son incorrectas, pues no se están considerando todos los casos posibles.

Análisis del ítem 11 del cuestionario 2

D5. Un investigador quiere relacionar el número de cilindros de una serie de marca de automóviles con el sitio donde se han fabricado. Obtiene las siguientes medidas de asociación en sus datos:

		Valor
Lambda	Número de cilindros dependiente	,183
Lambda	País de origen dependiente	,000
	V de Cramer	,445
	C Coeficiente de contingencia	,532

- Los valores de Lambda indican que se obtiene mejor predicción al predecir el número de cilindros sabiendo el país de origen que al predecir el país de origen sabiendo el número de cilindros.**
- Los valores de Lambda indican que se obtiene mejor predicción al predecir el país de origen sabiendo el número de cilindros que al predecir el número de cilindros sabiendo el país de origen.
- Se obtiene igual predicción en ambos casos

Con este ítem se analiza el conocimiento del estudiante de las medidas de asociación basadas en la reducción proporcional del error; para ser más concretos, sobre la Lambda de Goodman y Kruskal (Ruiz-Maya, Martín, Montero y Uriz, 1995). Para responder, el estudiante tiene que comprender el significado de dicho coeficiente, conocer su rango e incluso distinguir entre los dos posibles valores de Lambda, según la variable independiente sea la fila o la columna. La respuesta correcta es la a), pues la tabla proporciona un valor mayor (18,3%) de error reducido al predecir el valor de la variable “Número de cilindros”, sabiendo “País de origen” que al contrario.

- Si el alumno elige b) daría una respuesta falsa confundiendo los conceptos de variable dependiente e independiente, un error detectado por Ruiz (2006).
- Otro error posible es elegir c), pues los valores de Lambda que aparecen en la tabla son diferentes, por lo que las predicciones son diferentes según cual sea la variable dependiente.

Análisis del ítem 12 del cuestionario 2

D6. En el ejemplo anterior, los valores de los coeficientes de asociación V y C indican:

- a. Una asociación positiva débil
- b. Una asociación positiva moderada
- c. Una asociación negativa moderada
- d. Tendríamos que realizar el contraste Chi-cuadrado para conocer la intensidad y el signo

Con este ítem se quiere analizar si el estudiante comprende las medidas de asociación en tablas rxc, en particular, el coeficiente de contingencia de Pearson y el valor V de Cramer (Ruiz-Maya, Martín, Montero y Uriz, 1995). Para responder la pregunta, el estudiante tiene que conocer la interpretación de dichos coeficientes.

- Todas las respuestas son incorrectas pues estos coeficientes no informan del signo sino sólo de la intensidad de la asociación. En el ejemplo la asociación es moderada, pero no tiene sentido hablar del signo.
- Tampoco es necesario el contraste Chi-cuadrado para deducir la intensidad de la asociación; además, este coeficiente no nos informaría del signo (opción d).

5.3.3. ANÁLISIS DE LA PRUEBA FINAL

El último paso de la evaluación estuvo constituido por un conjunto de 4 ítems, cada uno de ellos con cuatro opciones, donde únicamente hay una opción correcta. Estos ítems formaron parte del examen final de la asignatura, donde hay en total 20 preguntas en el examen con la misma estructura. A continuación se describen y analizan estas preguntas, describiendo su contenido y los posibles errores que, a priori, es previsible que encontremos en las respuestas de los estudiantes.

Análisis del ítem 1 de la prueba final

E1. Se quiere estudiar la posible asociación entre si una madre gestante fuma durante su embarazo y si el niño es bajo de peso al nacer. En un estudio se obtienen los siguientes datos:

Gestantes	Recién nacidos de bajo peso	
	Sí	No
Fumadoras	43	207
No fumadoras	105	1645

¿Qué conclusiones puede obtener?

- Hay una relación directa entre las variables; se puede observar al comparar las celdas que están en las diagonales sin necesidad de calcular el estadístico Chi-cuadrado.**
- No hay relación pues la mayoría de las madres tiene un niño de peso normal y no fuma durante el embarazo.
- Hay relación inversa pues la mayoría de las madres fumadoras tiene un niño de peso normal y son muy pocas las madres fumadoras con niños bajos de peso.
- No cumple las propiedades de aplicación del estadístico Chi-cuadrado, por lo tanto no podemos dar ninguna conclusión referida a la dependencia.

Este ítem evalúa la comprensión del estudiante de procedimientos informales para evaluar la asociación, aplicados a un caso particular; así como su comprensión de la relación existente entre las celdas y la información que aportan los datos de la diagonal principal. Estas dos celdas (presencia-presencia y ausencia-ausencia) informarían que la asociación en la tabla es directa, según Inhelder y Piaget (1995). Además se incorpora la evaluación de las propiedades necesarias para la aplicación del estadístico Chi-cuadrado. La respuesta correcta es la a), pues en la diagonal principal están los valores que informan de dependencia directa, mientras que las otras dos celdas serían favorables a una asociación inversa. El resto son incorrectas, por las razones siguientes:

- En la respuesta b) solamente se está utilizando la celda de mayor frecuencia de la tabla, sin considerar el resto de celdas, lo que supondría una percepción incompleta de la asociación de acuerdo a Inhelder y Piaget (1955) y una concepción local de la asociación, según Estepa (1993). Además se aprecia que hay una asociación directa entre las variables, por estar en la diagonal principal una gran mayoría de sujetos del experimento.
- En la opción c) únicamente se usa la fila de madres fumadoras de la tabla, sin considerar las no fumadoras, mostrando una estrategia algo más desarrollada según Inhelder y Piaget (1955), pero que todavía corresponde a una concepción local de la asociación (Estepa, 1993). Además, como se ha dicho, hay asociación directa entre las variables.
- La respuesta d) es incorrecta, pues en este ejemplo se cumplen las propiedades de aplicación del estadístico Chi-cuadrado, ya que ninguna frecuencia es menor que 5.

El estudiante tiene que recordar estas condiciones y saber reconocerlas en un ejercicio.

Análisis del ítem 2 de la prueba final

- E2. Dos variables en una tabla de contingencia son independientes entre sí. Indica cuál de las siguientes propiedades NO SE CUMPLE
- Las frecuencias observadas de todas las celdas son iguales a las frecuencias teóricas esperadas.
 - En todas las celdas son iguales al producto del total por fila y columna que le corresponda, es decir $h_{i,j} = h_i \cdot h_j$
 - La suma de las frecuencias de las celdas en la diagonal principal será mayor que la suma de frecuencias en las celdas que no están en esta diagonal.**
 - Las frecuencias relativas condicionales por filas son iguales a las frecuencias relativas marginales por filas.

Este ítem evalúa la comprensión de la idea de independencia y el conocimiento de sus propiedades en una tabla de contingencia. La respuesta correcta es la c), pues se pregunta cual propiedad no se cumple, y esta sólo se cumple si las frecuencias marginales son iguales, pero no en el caso general (Batanero y Díaz, 2008). El resto de opciones son erróneas:

- En la opción a) no se tiene en cuenta que en caso de independencia, las frecuencias observadas coinciden con las frecuencias esperadas. Igualmente ocurre en la respuesta.
- La respuesta d) es incorrecta, ya que en caso de independencia estas distribuciones coinciden.

Análisis del ítem 3 de la prueba final

- E3. Un psicólogo pone a punto una terapia para disminuir la ansiedad. Recoge datos de 200 pacientes a 100 de los cuáles (elegidos al azar) aplica la terapia y al resto otra alternativa y pasados tres meses de tratamiento analiza la posible asociación entre terapia seguida (nueva o antigua) y disminución del nivel de ansiedad (disminuye o no disminuye). Aplica un contraste Chi-cuadrado de independencia y obtiene un valor $\chi^2=49,1$, con 1 g.l. ¿Qué conclusión debiera obtener de estos resultados?
- El resultado del contraste no es estadísticamente significativo, pues el valor en la tabla así lo confirma al comparar con el obtenido por el investigador. Debiera concluir que las variables no están relacionadas entre sí.
 - El resultado del contraste es estadísticamente significativo, pues $p < 0,05$. Debiera concluir que las variables están relacionadas entre sí. La relación es muy pequeña, porque el valor p es muy bajo.
 - El resultado del contraste es estadísticamente significativo, pues $p < 0,05$. Debiera concluir que las variables están relacionadas entre sí. Para saber si la relación es fuerte o débil tenemos que calcular alguna medida de asociación.**
 - No podemos concluir si hay relación entre las variables sólo con el valor Chi-cuadrado. Tenemos que calcular los coeficientes de asociación.

Este ítem evalúa la comprensión del estudiante del procedimiento de contraste de independencia y la interpretación de sus resultados. En el ejemplo aparecen el valor Chi, el valor p , y los grados de libertad, siendo el resultado estadísticamente significativo. Además se evalúan el conocimiento de los criterios de aceptación o rechazo de la hipótesis nula. Por último se incorpora una pregunta sobre los coeficientes de asociación para observar si el estudiante comprende su utilidad. La respuesta correcta es: la c), pues un valor p bajo (como es nuestro caso por ser Chi grande) indica asociación entre las variables, y además la forma de valorar la intensidad es mediante una medida de asociación (Batanero y Díaz, 2008).

- La opción a) es falsa, pues el resultado indica que el contraste es significativo, y, por lo tanto, las variables se deben considerar como dependientes. El estudiante podría confundir la región crítica y de aceptación, o bien, hacer un error al utilizar las tablas de la distribución Chi-cuadrado, que se les dejaron en el examen.
- La respuesta b) es incorrecta, ya que el hecho de que el Chi-cuadrado sea estadísticamente significativo no informa de la intensidad de la asociación, sino que requiere algún coeficiente. El estudiante podría confundir intensidad de la asociación con valor p .
- La respuesta d) es también errónea, pues con el valor Chi obtenido podemos rechazar la hipótesis de independencia, lo que es equivalente a asumir la asociación entre las variables.

Análisis del ítem 4 de la prueba final

E4. Un investigador quiere relacionar el número de cilindros de una serie de marca de automóviles con el sitio donde se han fabricado (Europa, Estados Unidos o Japón). Obtiene las siguientes medidas de asociación en sus datos: V de Cramer= 0,445; Coeficiente de contingencia de Pearson= 0,532; Lambda de Goodman y Kruskal (Variable dependiente= Número de cilindros)= 0,183.

¿Qué conclusión debe obtener el investigador?

- Hay una relación muy alta entre las variables y se puede predecir muy fácilmente el número de cilindros del coche, si se sabe donde se ha fabricado
- Hay una relación moderada-alta entre las variables y se puede predecir muy fácilmente el número de cilindros del coche, si se sabe donde se ha fabricado
- Hay una relación moderada-alta entre las variables, pero no es fácil predecir el número de cilindros del coche, si se sabe donde se ha fabricado**
- Hay una relación muy alta entre las variables, pero no es fácil predecir el número de cilindros del coche, si se sabe donde se ha fabricado

Este ítem evalúa el conocimiento del estudiante de los coeficientes de asociación en tablas rxc y de los coeficientes basados en la reducción proporcional del error (tomado de Ruiz-Maya, Martín, Montero y Uriz, 1995). La respuesta correcta es la c), ya que los coeficientes de asociación V y C informan que la relación es moderada-alta (pues su valor se sitúa en el medio de su posible rango), y el coeficiente Lambda nos indica que no es fácil predecir el número de cilindros, si se conoce donde se ha fabricado, pues su valor es pequeño, por lo que la reducción obtenida del error es sólo del 18%. El resto de opciones son incorrectas:

- La respuesta a) porque la relación no es muy alta, debido a que los valores de los coeficientes V de Cramer y de contingencia son moderados. Además, como Lambda de Goodman y Kruskal es muy bajo no es fácil la predicción.
- La respuesta b) es incorrecta, ya que Lambda de Goodman y Kruskal es muy bajo, por lo que no es fácil la predicción, aunque la primera parte del enunciado es cierta.
- En la respuesta d) se cumple la primera parte del enunciado.

5.4. SINTESIS DE LAS EVALUACIONES MEDIANTE CUESTIONARIO

Como síntesis del análisis realizado, mostramos en la Tabla 5.4.1 los objetos matemáticos incluidos en los ítems usados en el estudio (pruebas tipo test y preguntas del examen final). Observamos que se han tenido en cuenta la mayor parte de los objetos incluidos en la enseñanza, lo que dotará de validez de contenido al proceso de evaluación. En nuestro marco teórico podemos decir que el cuestionario tiene una alta idoneidad epistémica, pues su contenido es representativo del significado institucional implementado en la enseñanza.

Tabla 5.4.1. Objetos considerados en la autoevaluación y prueba final

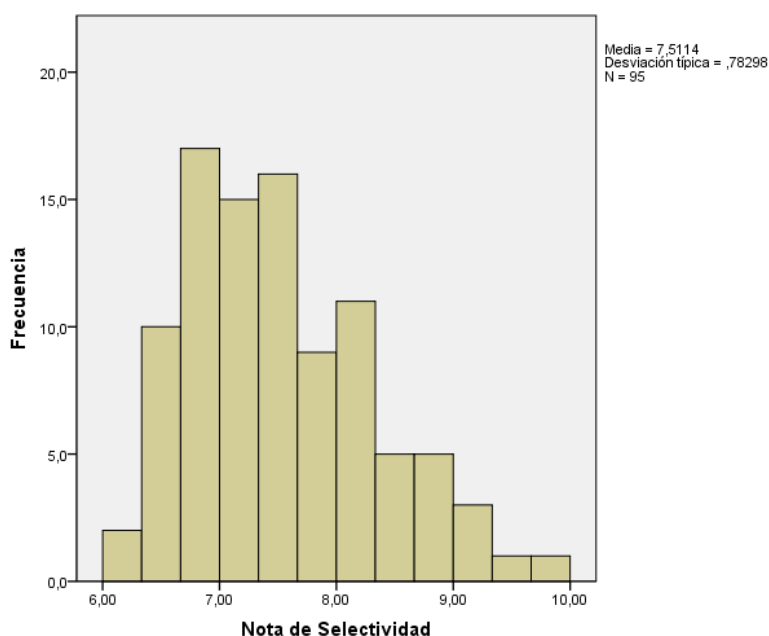
	Objetos matemáticos	Ítem
Situaciones-problema	P1. Resumir la información obtenida de un conjunto de observaciones bivariantes de modo que se visualicen posibles relaciones entre las variables.	A6,E3,E4
	P2. Cálculo de probabilidades asociadas a valores específicos de una o las dos variables.	A2,A3,A4
	P3. Contraste de homogeneidad.	C4
	P4. Contraste de independencia.	C4,E3
	P5. Definir una medida de la intensidad de asociación entre las variables.	D1,D2,D3,D4,D6,E4
Lenguaje	R1. Diagrama de barras adosado.	A1,A5
	R2. Diagrama de barras apilado.	A1,A5
	R3. Gráfico tridimensional.	A1,A5
	R4. Lenguaje verbal.	Todas
	R5. Lenguaje simbólico.	B3,B5,B6,C1,C4,C5,D4,E2,E3
	R6. Lenguaje gráfico.	A1,A5
	R7. Tablas (de frecuencia y de contingencia).	B4,C5,D5,E1
Procedimientos	PR1. Lectura e interpretación de la tabla.	B4,C5,D5,E1
	PR2. Construcción de una tabla de contingencia.	A6
	PR3. Cálculo de probabilidades simples, compuestas y condicionales a partir de la tabla.	A2,A3,A4
	PR4. Cálculo de frecuencias condicionales y marginales (absolutas o relativas).	A2,A3,A4
	PR5. Representación gráfica de datos a partir de una tabla de contingencia.	A1,A5
	PR6. Cálculo de las frecuencias esperadas.	B2
	PR7. Comprobar las propiedades de independencia.	B1,E1,E2
	PR8. Selección de medidas de la asociación.	D2,E4
	PR9. Cálculo e interpretación de medidas en tablas 2x2.	D1,D3
	PR10. Cálculo e interpretación de medidas en tablas rxc.	D6,E4
	PR11. Cálculo e interpretación de medidas basadas en la reducción del error.	D5,E4
	PR12. Formulación de hipótesis (contraste de independencia).	C4,E3
	PR13. Formulación de hipótesis (contraste de homogeneidad).	C4
	PR14. Interpretación de resultados del test Chi-cuadrado obtenido con el ordenador en el contexto del problema.	C2,C3,C4,C5,E3
	PR15. Completar un contraste de hipótesis tomando una decisión.	C4,C6,E3
	PR16. Manejo del software.	C5
	PR17. Cálculo de los grados de libertad para realizar un test Chi-cuadrado.	C1
	PR18. Manejo de tablas de distribuciones de probabilidad.	C2,E3
	PR19. Cálculo del estadístico de contraste.	C2
Conceptos	C1. Variable simple / doble valor.	Todas
	C2. Distribución bidimensional.	Todas
	C3. Variable cualitativa y cuantitativa.	A6
	C4. Tablas de contingencia, tablas rxc.	A5,E3
	C5. Frecuencia absoluta y relativa doble.	A1,A3,A4,B2,B3,E1,E2
	C6. Frecuencia absoluta y relativa marginal; distribución marginal.	A1,A2,A4,B1,B2,B3,E2
	C7. Frecuencia relativa condicional; distribución condicional.	A1,A4,B1,E1,E4
	C8. Probabilidad simple.	A2,A3
	C9. Probabilidad conjunta y condicional.	A3,A4
	C10. Independencia /Asociación, dependencia funcional.	B1,B4,B6,C3,C4,C6,E1,E3
	C11. Asociación inversa, directa.	B4,E1
	C12. Intensidad de la asociación.	D1,D5,D6,E3,E4
	C13. Causalidad.	B5

	Objetos matemáticos	Ítem
	C14. Diagrama de barras apilado.	A1,A5
	C15. Diagrama de barras adosado.	A1,A5
	C16. Gráfico tridimensional.	A5
	C17. Frecuencias esperadas.	B2,E2
	C18. Estadístico Chi-cuadrado, grados de libertad, distribución.	C1,C2,C3,C4,C5, C6,E3
	C19. Coeficiente Phi, Riesgo relativo, Razón de productos cruzados (medidas tablas 2x2).	D1,D2,D3,D4
	C20. Coeficiente C y V (medidas simétricas, tablas $r \times c$).	D2,D6,E4
	C21. Coeficiente Lambda (medidas asimétricas).	D5,E4
	C22. Significación.	C2,C4,C5,E3
	C23. p valor.	C3,C4,C5,C6,E3
	C24. Contraste de hipótesis. Hipótesis alternativa y nula.	C4
	C25. Población y muestra. Estadístico de contraste.	C1,C2,C3,C4,C5, C6,E3
	C26. Región crítica y de aceptación.	C1,C2,C3,C4,C5, C6,E3
	C27. Posibles errores en un contraste de hipótesis.	C1,C2,C3,C4,C5, C6,E3
Propiedades	PP1. Relaciones entre frecuencias absolutas y relativas.	A2,A3
	PP2. Relaciones entre frecuencias dobles, condicionales y marginales.	A3,A4,E2
	PP3. Igualdad /diferencia de frecuencias relativas condicionales en caso de independencia/dependencia.	B1,E1,E2
	PP4. Igualdad/ diferencia de frecuencias relativas condicionales y marginales en caso de independencia/ dependencia.	B1,E1,E3
	PP5. Celdas que indican signo de asociación en la tabla 2x2.	B4,E1,E2
	PP6. Suma de frecuencias relativas igual a la unidad.	A2
	PP7. Frecuencias esperadas en caso de independencia.	B2,E2
	PP8. Las hipótesis de un contraste son complementarias.	C1,C2,C3,C4,C5, C6,E3
	PP9. Relación entre asociación y causalidad.	B5
	PP10. Diferencia entre dependencia aleatoria y funcional.	B6
	PP11. Relación entre regiones de aceptación y rechazo y decisión en el contraste.	C2,C3,C4,C5,E3
	PP12. Relación entre valor p en el contraste Chi-cuadrado y asociación.	C3,E3
	PP13. Diferencias entre el significado y aplicación de las medidas de asociación.	D2,E4
	PP14. El diagrama de barras adosado se puede construir con frecuencias absolutas o porcentajes.	A1
	PP15. Relación y diferencia entre el diagrama de barras adosado y apilado.	A1,A5
	PP16. Propiedades de independencia.	B3,E2
	PP17. Diferenciar valores de Chi-cuadrado estadísticamente significativos.	C2,C5
	PP18. Relaciones entre los errores en un contraste.	C1,C2,C3,C4,C5, C6,E3
	PP19. Condiciones de aplicación del estadístico Chi-cuadrado.	C5,E1
	PP20. Relación entre el riesgo relativo por filas y columnas.	D3
	PP21. Rango de acción de las medidas de asociación.	D4,D5,D6,E4
	PP22. Relación entre probabilidad y tabla de contingencia.	A2,A3,A4
	PP23. Suma de frecuencias absolutas dobles igual al total de la muestra.	A2,A3
Argu	A1. Razonamientos informales.	B4,E1
	A2. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar.	C5,C6,D1,D5,D6
	A3. Deducción de propiedades sin excesiva formalización.	B4,E1

5.5. RESULTADOS EN EL CUESTIONARIO

En la enseñanza participaron 104 estudiantes de Psicología, de ellos 51 estudiantes en el primer grupo (8 hombres y 43 mujeres), y 53 estudiantes en el segundo grupo (16 hombres y 37 mujeres). Todos ellos eran estudiantes de segundo año y cursaban la asignatura *Técnicas de análisis en la investigación psicológica* en la Universidad de Granada. La mayoría tenía una edad de 19 o 20 años.

Figura 5.5.1. Histograma de las calificaciones de Selectividad



En la Figura 5.5.1 observamos que los estudiantes tenían una nota de acceso en la Universidad variada, pero con una media superior (7,51) a la obtenida en el estudio del Capítulo 3 (donde la media resultó ser de 6,72). Además, estos alumnos venían principalmente del bachillerato de Ciencias Sociales, al igual que en la muestra del Capítulo 3 (Tabla 5.5.1). Por otro lado, sólo 4 estudiantes eran repetidores y ninguno había estudiado el tema con anterioridad.

Tabla 5.5.1. Frecuencia y porcentaje de la especialidad de Bachillerato en los estudiantes de la muestra

	Frecuencia	Porcentaje
Ciencias de la Salud	17	16,35
Ciencias Sociales	64	61,54
Tecnológico	10	9,62
Ciclo Superior	2	1,92
Arte	3	2,88
No contesta	8	7,69
Total	104	100

Los cuestionarios fueron cumplimentados individualmente y sin ayuda por los estudiantes, cada uno de ellos en una sesión de clase de 1 hora de duración. Se recomendó a los estudiantes que completasen las preguntas con interés, motivándoles a colaborar, tanto porque se tuvo en cuenta la actividad en la calificación de la asignatura, como porque se les indicó que los resultados se utilizarían en la mejora del tema en cursos sucesivos. El investigador, así como un observador participante y el profesor de la asignatura estuvieron presentes y, en caso de duda sobre el significado de alguna frase, se clarificó dicho significado al estudiante, aunque las preguntas en este sentido fueron muy pocas.

Las cuatro preguntas incluidas en la prueba final de la asignatura también se cumplimentaron individualmente, en la misma forma que el resto de dicha prueba final. En este caso, el profesor de la asignatura fue el que atendió las posibles preguntas de los estudiantes y el investigador asistió simplemente como observador participante.

A continuación se analizan los resultados obtenidos en ambos cuestionarios y en la prueba final. El número de participantes varía de uno a otro ítem, pues al haberse incorporado la enseñanza al curso ordinario que seguían estos estudiantes no suele ser obligatoria la asistencia. Aún así, como veremos hay una fuerte participación, lo que indica que se tuvo éxito al interesar a los estudiantes por este tema.

5.5.1. RESULTADOS EN EL PRIMER CUESTIONARIO

El primer cuestionario, correspondiente a los dos primeros temas de la enseñanza, lo completaron 94 estudiantes. A continuación se presentan y analizan para cada ítem las respuestas correctas, incorrectas y no respuestas en cada apartado (marcados como apartados a, b y c).

Resultados del primer ítem del cuestionario 1

A1. La principal diferencia entre el diagrama de barras apilado y el adosado es:

- a. El diagrama de barras apilado sólo se puede construir con frecuencias absolutas, mientras que el adosado se puede construir con frecuencias absolutas o porcentajes
- b. En el diagrama de barras apilado las frecuencias condicionales de cada valor de X para un mismo valor de Y se representan en una misma barra y en el adosado se representan una al lado de la otra.**
- c. En el diagrama de barras apilado las frecuencias marginales de cada valor de X se representan en una misma barra y en el adosado se representan una al lado de la otra

En esta pregunta todos los apartados fueron respondidos correctamente por más de un 45% de la muestra, y un total de 28 estudiantes responden correctamente los tres apartados, comprendiendo perfectamente la finalidad de cada gráfico, 56 responden incorrectamente al menos a un apartado y 10 dejan algún apartado en blanco.

Tabla 5.5.1.1. Frecuencias (y porcentajes) de respuestas en el ítem A1 (n=94)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	71 (75,5)	14 (14,9)	9 (9,6)
b (Verdadera)	42 (44,7)	44 (46,8)	8 (8,5)
c (Falsa)	50 (53,2)	35 (37,2)	9 (9,6)

Lo más sencillo en este ítem fue reconocer que se puede usar diferentes tipos de frecuencias en los gráficos (opción a), resultando más difícil diferenciar qué tipo de información se representa en las barras adosadas o apiladas (opciones b y c). Algunos de estos estudiantes podrían no alcanzar el nivel de *extracción de datos*, que consiste en poner en relación un elemento de un eje de un gráfico con el de otro eje, pues no llegarían a comprender qué tipos de frecuencia se representa en cada gráfico para cada valor de la variable (Bertin, 1967).

Resultados del ítem 2 del cuestionario 1

A2. En una tabla de contingencia, la suma de las frecuencias relativas marginales por filas es:

- a. **Igual a la suma de frecuencias relativas marginales por columnas**
- b. Igual al total de la muestra
- c. **Igual a la suma de las frecuencias absolutas marginales dividida por el total de la muestra**

Tabla 5.5.1.2. Frecuencias (y porcentajes) de respuestas en el ítem A2 (n=94)

Respuesta	Correcta	Incorrecta	No responde
a (Verdadera)	53 (56,4)	41 (43,6)	0 (0)
b (Falsa)	61 (64,9)	33 (35,1)	0 (0)
c (Verdadera)	37 (39,4)	57 (60,6)	0 (0)

En este ítem, donde se relacionan las frecuencias relativas marginales con otros elementos, una alta proporción de estudiantes reconoce que la suma de las frecuencias relativas marginales por filas no es igual al total de la muestra, y si es igual a las frecuencias relativas marginales por columna (opciones a y b). Fué más difícil ver que estas frecuencias son iguales a la suma de las frecuencias absolutas marginales dividida por el total de la muestra (opciones c) (Nortes, 1993).

En esta pregunta un total de 15 estudiantes responden correctamente los tres apartados (16%) y 79 estudiantes tienen al menos un apartado con respuesta incorrecta

(84%), en general en la opción (c), donde tienen que relacionar formulas; es posible que algunos de estos estudiantes tienen dificultades con la simbolización algebraica.

Resultados del ítem 3 del cuestionario 1

A3. Las frecuencias dobles absolutas y relativas están relacionadas entre sí mediante:
a. Las frecuencias condicionales.
b. Las frecuencias marginales.
c. El tamaño de la muestra.

Tabla 5.5.1.3. Frecuencias (y porcentajes) de respuestas en el ítem A3 (n=94)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	76 (80,9)	15 (16)	3 (3,2)
b (Falsa)	73 (77,7)	18 (19,1)	3 (3,2)
c (Verdadera)	65 (69,1)	26 (27,7)	3 (3,2)

En este ítem se obtuvieron muy buenos resultados en todos sus apartados, de lo que se deduce que los estudiantes adquirieron la relación entre las frecuencias absolutas y relativas y no confunden estos conceptos. Lo más sencillo fue recordar que no se pueden relacionar las frecuencias dobles absolutas y relativas mediante las frecuencias condicionales (opción a; Nortes, 1993); seguido por reconocer que no se relacionan mediante las frecuencias marginales pero si por el tamaño de la muestra (opciones b y c). Un total de 57 (60,64%) estudiantes responden correctamente los tres apartados; este alto porcentaje nos lleva a pensar que la enseñanza ha permitido establecer la relación entre las frecuencias absolutas y relativas de forma bastante favorable. Por otro lado, 34 tienen al menos un apartado incorrecto y sólo 3 dejan algún apartado en blanco.

Resultados del ítem 4 del cuestionario 1

A4. Para calcular las frecuencias relativas condicionadas es necesario:
a. Sólo las frecuencias absolutas dobles de la fila o columna por la que se condiciona.
b. Las frecuencias relativas de la fila o columna por la que se condiciona y el total de la muestra.
c. Las frecuencias absolutas o relativas de la fila o columna por la que se condiciona.

Tabla 5.5.1.4. Frecuencias (y porcentajes) de respuestas en el ítem A4 (n=94)

Respuesta	Correcta	Incorrecta	No responde
a (Verdadera)	17 (18,1)	74 (78,7)	3 (3,2)
b (Falsa)	50 (53,2)	40 (42,6)	4 (4,3)
c (Verdadera)	34 (36,2)	56 (59,6)	4 (4,3)

Este ítem sobre el cálculo de las frecuencias relativas condicionales, da resultados contradictorios, con bastante diferencia de dificultad, según el apartado. Los

alumnos comprenden que para el cálculo de las frecuencias condicionales no se requiere el total de la muestra (opción b); fue más difícil reconocer que pueden calcularse mediante las frecuencias absolutas dobles o relativas de la fila o columna por la que se condiciona (opciones a y c; Amon, 1993).

En esta pregunta sólo 6 estudiantes responden correctamente los tres apartados (6,4%), lo cual indica su dificultad; 84 tienen al menos un apartado incorrecto (89,4%) y 4 dejan alguno en blanco (4,3%). Los estudiantes aprendieron la fórmula tal como se les dio pero tuvieron poca facilidad para deducir fórmulas alternativas.

Resultados del ítem 5 del cuestionario 1

- A5. Las tablas rxc se pueden representar gráficamente en los gráficos:
- Únicamente en el diagrama de barras apilado y el adosado.
 - Únicamente en el gráfico tridimensional y el diagrama de barras adosado.
 - En los mismos gráficos que una tabla 2x2.**

Tabla 5.5.1.5. Frecuencias (y porcentajes) de respuestas en el ítem A5 (n=94)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	83 (88,3)	3 (2,9)	8 (8,5)
b (Falsa)	64 (68,1)	22 (23,4)	8 (8,5)
c (Verdadera)	59 (62,8)	27 (28,7)	8 (8,5)

Este ítem sobre gráficos resultó muy sencillo, sobre todo reconocer que estas tablas se pueden representar en gráficos diferentes del diagrama de barras apilado y el adosado (opción a); seguido por recordar gráficos aparte del tridimensional y el diagrama de barras adosado y reconocer que se pueden representar en los mismos gráficos que una tabla 2x2 (opciones b y c). Hay un porcentaje bastante alto de aciertos, pues 56 estudiantes responden correctamente los tres apartados (59,6%), lo que indica que han comprendido la relación entre el tipo de tabla y los gráficos; 30 tienen al menos un error (31,9%) y 8 dejan algún apartado en blanco (8,5%).

Análisis del ítem 6 del cuestionario 1

- A6. El objetivo de construir una tabla de contingencia es:
- Resumir los datos de dos variables cuantitativas
 - Resumir los datos de dos variables cualitativas**
 - Resumir los datos de dos variables cualitativas, o bien numéricas con pocos valores diferentes**

Tabla 5.5.1.6. Frecuencias (y porcentajes) de respuestas en el ítem A6 (n=94)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	53 (56,4)	40 (42,6)	1 (1,1)
b (Verdadera)	51 (54,3)	42 (44,7)	1 (1,1)
c (Verdadera)	18 (19,1)	75 (79,8)	1 (1,1)

Los resultados son intermedios, por lo que no es claro si los estudiantes comprenden los objetivos en la construcción de una tabla de contingencia. Lo más sencillo fue recordar que no se utilizan variables cuantitativas en general, sino sólo si tienen pocos valores, en la construcción de tablas de contingencia. En esta pregunta un total de 5 estudiantes responden correctamente los tres apartados (5,3%), 88 tienen al menos un fallo (93,6%) y 1 dejan algún apartado en blanco (1,1%).

Resultados del ítem 7 del cuestionario 1

- B1. 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.**

Tabla 5.5.1.7. Frecuencias (y porcentajes) de respuestas en el ítem B1 (n=94)

Respuesta	Correcta	Incorrecta	No responde
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)

Este es el primero de una serie de ítems sobre comprensión de las propiedades de la independencia. Los estudiantes parecen comprender que para que dos variables de una tabla de contingencia sean independientes, las frecuencias relativas condicionales por columna y por fila (opciones a y b) han de ser iguales. En principio, esto supone que los estudiantes están razonando en el nivel 5 de Pérez Echeverría y, en consecuencia, es previsible que en la resolución de juicios de asociación usen estrategias de nivel bastante más alto que las empleadas en el estudio de evaluación inicial. Ha sido más difícil reconocer que también tienen que ser iguales las frecuencias relativas condicionales y frecuencias relativas marginales (opción c; Batanero y Díaz, 2008), donde hubo un pequeño porcentaje de aciertos; posiblemente se hizo menos énfasis en esta propiedad en la enseñanza.

En esta pregunta sólo 8 estudiantes responden correctamente los tres apartados (8,5%), 83 tienen al menos un error (88,3%) y 3 dejan algún apartado en blanco (3,2%), viéndose que la opción (c) es donde se cometen la mayoría de los errores. Es posible

que los estudiantes esperasen la palabra “correspondiente” al final de la frase, que se ha dejado implícita; pues de hecho las que han de ser iguales son las frecuencias condicionales y la marginal correspondiente.

Resultados del ítem 8 del cuestionario 1

- B2. Las frecuencias esperadas se calculan mediante:
- Las frecuencias absolutas dobles y el total de la muestra.
 - Las frecuencias relativas dobles y el total de la muestra.
 - Las frecuencias absolutas marginales y el total de la muestra.**
 - Las frecuencias relativas marginales

Tabla 5.5.1.8. Frecuencias (y porcentajes) de respuestas en el ítem B2 (n=94)

Respuesta	Correcta	Incorrecta	No responde
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)

El ítem en general fue sencillo; pues los estudiantes reconocen las fórmulas erróneas de cálculo de las frecuencias esperadas (opciones b y d); fue algo más difícil recordar que estas frecuencias se calculan también mediante las frecuencias marginales (opción c), posiblemente porque los estudiantes no llegaron a realizar personalmente los cálculos, salvo en ocasiones aisladas, sino que utilizaron la hoja Excel. En total 55 estudiantes responden correctamente los tres apartados (58,5%), 37 tienen al menos una respuesta errónea (39,4%) y 2 dejan algún apartado en blanco (2,1%).

Resultados del ítem 9 del cuestionario 1

- B3. En caso de que haya asociación entre variables, las frecuencias relativas dobles:
- En todas las celdas son iguales al producto del total por fila y columna que le corresponda, es decir $h_{i,j} = h_i \cdot h_j$
 - Puede ocurrir que coincida en alguna celda al producto del total por fila y columna que le corresponda, es decir $h_{i,j} = h_i \cdot h_j$**
 - Nunca son iguales al producto del total por fila y columna que le corresponda, es decir nunca se cumple $h_{i,j} = h_i \cdot h_j$

Tabla 5.5.1.9. Frecuencias (y porcentajes) de respuestas en el ítem B3 (n=94)

Respuesta	Correcta	Incorrecta	No responde
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)

En este ítem se trata de evaluar si los alumnos comprenden una de las propiedades de independencia. Los alumnos no recordaron que en caso de asociación podrían ser iguales los productos de las frecuencias relativas marginales por fila y columna que le corresponda y la frecuencia esperada, es decir, podría darse $h_{ij} = h_{i.} \cdot h_{.j}$ en alguna celda (opción c) (Batanero y Díaz, 2008); paradójicamente rechazan la forma afirmativa de la misma proposición (opción b). Deducimos que los alumnos tienen dificultades de lectura del ítem, ya que también la opción (a) se toma por correcta en una gran parte de la muestra.

En esta pregunta un total de 10 estudiantes responden correctamente los tres apartados (10,6%), 80 tienen al menos un apartado mal (85,1%) y 4 dejan algún apartado en blanco (4,3%). En este ítem profundizamos en una de las propiedades de independencia, tratando de ver si los alumnos la aplican al caso de asociación, lo que no ha sido sencillo. El simbolismo usado también puede haber dificultado la respuesta.

Resultados del ítem 10 del cuestionario 1

B4. 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

En este ejercicio, los tres apartados han resultado sencillos, reconociendo los estudiantes que las celdas de la diagonal principal, es decir, las celdas correspondientes a presencia de los dos caracteres o ausencia de los dos caracteres, informan de una dependencia directa; y las celdas con presencia de uno de los dos caracteres y ausencia del otro, informan de una dependencia inversa (Inhelder y Piaget, 1955).

Tabla 5.5.1.10. Frecuencias (y porcentajes) de respuestas en el ítem B4 (n=94)

Respuesta	Correcta	Incorrecta	No responde
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)

Un total de 89 (94,7%) estudiantes responden correctamente los tres apartados, lo que indica una comprensión general de esta propiedad por los estudiantes. Es de esperar que ello favorezca la superación las estrategias de bajo nivel (Pérez Echeverría, 1990) como es fijarse en la primera celda (c) o en una única distribución condicionada (a) al realizar juicios de asociación y que aparecieron con frecuencia apreciable en el estudio inicial de evaluación (Capítulo 3). Sólo 3 estudiantes tienen al menos un apartado incorrecto (3,2%) y 2 dejan alguno en blanco (2,1%), lo que nos da una indicación de que la enseñanza recibida sirvió para mejorar las estrategias intuitivas de los estudiantes.

Análisis del ítem 11 del cuestionario 1

B5. 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.		

Tabla 5.5.1.11. Frecuencias (y porcentajes) de respuestas en el ítem B5 (n=94)

Respuesta	Correcta	Incorrecta	No responde
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)

Este ítem, en el cual se relacionan los conceptos de asociación y causalidad, ha dado buenos resultados. Los estudiantes reconocieron que si hay relación causal entre A y B , entonces la asociación entre A y B no tiene por qué ser siempre positiva (opción a). Tampoco tuvieron dificultad para indicar que no siempre una asociación entre A y B se traduce una relación causal y que si hay una relación causal entre A y B , entonces habrá asociación, que puede ser positiva o negativa (Cook y Campbell, 1979; opciones b y c).

Un total de 72 estudiantes responden correctamente los tres apartados (76,6%), por lo que la mayoría aparentemente ha superado la concepción causal descrita por Estepa (1993), comprendiendo la relación entre correlación y causalidad, aunque en la experiencia realizada por dicho autor, sus estudiantes no siempre la superaron; 21 tienen al menos un fallo (22,3%) y 1 dejan algún apartado en blanco (1,1%).

Resultados del ítem 12 del cuestionario 1

B6. 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

En este ítem se analiza la comprensión de los estudiantes de la diferencia entre dependencia funcional y dependencia aleatoria. Prácticamente todos los estudiantes comprenden que en la dependencia aleatoria, a cada valor de la variable independiente X le corresponde más de un valor de la variable dependiente Y (opción a). También resultó razonablemente sencillo reconocer que tanto la dependencia aleatoria como la funcional pueden ser directa o inversa (opción c). Hubo mayor número de errores para recordar que, en la dependencia aleatoria, al variar X suele variar Y , pero no siempre (Nortes, 1993; opción b); Son pocos (35 estudiantes; 37,2%) los que responden correctamente los tres apartados, 54 tienen al menos uno incorrecto (57,5%) y 5 dejan algún apartado en blanco (5,3%).

Tabla 5.5.1.12. Frecuencias (y porcentajes) de respuestas en el ítem B6 ($n=94$)

Respuesta	Correcta	Incorrecta	No responde
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 (5,3)

5.5.2. RESULTADOS DEL SEGUNDO CUESTIONARIO

El segundo cuestionario incluye ítems sobre los temas 3 y 4 de la enseñanza, es decir, a la realización del contraste Chi-cuadrado y la comprensión de las medidas de asociación y fue completado por 93 estudiantes. A continuación se comentan los resultados.

Resultados del ítem 1 del cuestionario 2

C1. Para realizar un test Chi-cuadrado en una tabla de contingencia con m filas y n columnas, los grados de libertad del estadístico Chi-cuadrado son:	
a.	$m \times n$.
b.	$m + n$.
c.	$(m-1) \times (n-1)$.
d.	$(m-1) + (n-1)$.

Tabla 5.5.2.1. Frecuencias (y porcentajes) de respuestas en el ítem C1 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	93 (100)	0 (0)	0 (0)
b (Falsa)	93 (100)	0 (0)	0 (0)
c (Verdadera)	86 (92,5)	7 (7,5)	0 (0)
d (Falsa)	87 (93,5)	6 (6,5)	0 (0)

Este ítem, que evalúa la comprensión del cálculo de los grados de libertad fue muy sencillo, pues los estudiantes recuerdan que en dicho cálculo hay que restar uno al número de filas y al número de columnas (Martín y Luna, 1994), (opciones a y b); hubo algún fallo al diferenciar si se multiplica o se suman los valores, después de restarles uno (opciones c y d). Ello sugiere un aprendizaje memorístico de la fórmula de cálculo, ya que, si lo que vamos a calcular es el número de cantidades independientes en cada variable, fijada una, las demás son independientes y en total habrá que multiplicar $m-1$ por $n-1$ en lugar de sumarlas.

Un total de 86 estudiantes responden correctamente los tres apartados (92,5%) y 7 tienen al menos un apartado mal contestado (7,5%). En resumen, estos estudiantes han aprendido la fórmula de cálculo de los grados de libertad, a pesar que no se les exigía. Apenas aparece el error citado por Alvarado (2007) y Olivo (2008) consistente en no restar el número de restricciones en el cálculo de los grados de libertad.

Resultados del ítem 2 del cuestionario 2

- C2. Los grados de libertad sirven para:
- Contar el número de filas y columnas.
 - Calcular el parámetro de la distribución del estadístico.**
 - Encontrar en las tablas de la distribución la probabilidad de un valor de Chi-cuadrado, y saber si el valor de Chi cuadrado es estadísticamente significativo.**

Tabla 5.5.2.2. Frecuencias (y porcentajes) de respuestas en el ítem C2 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	90 (96,8)	3 (3,2)	0 (0)
b (Verdadera)	34 (36,6)	59 (63,4)	0 (0)
c (Verdadera)	87 (93,5)	6 (6,5)	0 (0)

Este ítem sobre la utilidad de los grados de libertad proporciona resultados mezclados. Fue sencillo para los estudiantes comprender que los grados de libertad no sirven para contar el número de filas y columnas; que los grados de libertad determinan la distribución de la probabilidad, y permiten hallar un valor crítico para decidir si el valor de Chi-cuadrado es estadísticamente significativo (Martín y Luna, 1994) (opciones a y c). Fue más difícil reconocer el término *parámetro* (opción b);

posiblemente por ser poco familiar al estudiante por no haber hecho suficiente énfasis en este punto. Un total de 28 estudiantes responden correctamente los tres apartados (30,1%) y 65 tienen al menos un apartado mal (69,9%).

Resultados del ítem 3 del cuestionario 2

C3. Un investigador usa un test Chi-cuadrado para determinar si hay asociación entre dos variables en una tabla de contingencia. Calcula el valor de Chi-cuadrado y obtiene el valor p correspondiente. ¿Cuál de los siguientes valores p indicaría de forma más concluyente que hay asociación entre las variables.

- | | |
|----------------|--------|
| a. .002 | b. .01 |
| c. .05 | d. .10 |

Tabla 5.5.2.3. Frecuencias (y porcentajes) de respuestas en el ítem C3 ($n=93$)

Respuesta	Correcta	Incorrecta	No responde
a (Verdadera)	38 (40,9)	55 (59,1)	0 (0)
b (Falsa)	81 (87,1)	12 (12,9)	0 (0)
c (Falsa)	75 (80,6)	18 (19,4)	0 (0)
d (Falsa)	56 (60,2)	37 (39,8)	0 (0)

En este ítem pide la interpretación del valor de p en un contraste de asociación (Lipschutz y Schiller, 1999). Los resultados son mezclados, pues aunque los estudiantes responden que los valores 0,01, 0,05 y 0,10 no son los más significativos (opciones b, c y d), les resultó más difícil reconocer que el más significativo de todos es el valor 0,002 (opción a).

En esta pregunta 33 estudiantes responden correctamente los tres apartados (35,5%) y 60 tienen al menos un error (64,5%). Pensamos que los fallos se producen por confundir p -valor e intensidad de asociación o bien por pensar que han de elegir uno de los valores 0,01 o 0,05 con los que los estudiantes están acostumbrados a trabajar (opciones b y d), ya que son los que más aparecen en la literatura y en los ejemplos y ejercicios trabajados en clase. En todo caso la dificultad de interpretación del p -valor ha sido citada por varios autores, como Vallecillos (1996) y Batanero (2000).

Resultados del ítem 4 del cuestionario 2

C4. Un investigador quiere estudiar si existe relación entre dos variables. Calcula el estadístico Chi-cuadrado y obtiene un valor p estadísticamente significativo ($p < 0,05$). La decisión que debe tomar es:

- a. **Rechazar la hipótesis de independencia entre las dos variables**
- b. Rechazar la hipótesis de asociación entre las dos variables
- c. Rechazar la hipótesis de asociación directa entre las dos variables
- d. Con estos datos, no podemos saber la decisión que se debe tomar

Tabla 5.5.2.4. Frecuencias (y porcentajes) de respuestas en el ítem C4 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Verdadera)	50 (53,8)	42 (45,2)	1 (1,1)
b (Falsa)	72 (77,4)	20 (21,5)	1 (1,1)
c (Falsa)	84 (90,3)	8 (8,6)	1 (1,1)
d (Falsa)	73 (78,5)	19 (18,3)	1 (1,1)

En este ítem se valora la comprensión de la lógica del contraste de hipótesis; más específicamente el conocimiento de la regla de decisión para aceptar o no la hipótesis nula en función de un valor p (Lipschutz y Schiller, 1999). Se ha formulado la decisión de varios modos, para ver si el estudiante discrimina las hipótesis nula y alternativa en este contraste. Los estudiantes recuerdan que con el valor obtenido no se puede rechazar la hipótesis de asociación o asociación directa entre las variables (opciones b y c), pues, por un lado, estas no son hipótesis nulas y por tanto, no se pueden rechazar. Por otro, el valor obtenido corrobora preferentemente la asociación y no la independencia.

Los alumnos también rechazan la idea de que con estos datos no se puede tomar una decisión (opción d), es decir, comprenden el significado del término “estadísticamente significativo”. Aunque con algo de menor frecuencia, más de la mitad de la clase toma la decisión correcta; rechazar la hipótesis de independencia entre las dos variables (opción a).

En esta pregunta un total de 49 estudiantes responden correctamente los tres apartados (52,7%), 43 tienen al menos un apartado mal (46,2%) y 1 dejan algún apartado en blanco (1,1%). Vemos como más de la mitad interpretan bien un resultado estadísticamente significativo, recordando correctamente las hipótesis de un contraste de independencia. Además es interesante el hecho de que muy pocos fallan en la opción (c), lo que hace pensar que se recuerda que en este contraste no se informa si la asociación es directa o inversa. Los fallos posiblemente sean debidos a que una parte de los estudiantes no comprende bien la diferencia entre hipótesis nula y alternativa, un punto difícil según Batanero (2000).

Resultados del ítem 5 del cuestionario 2

C5. Un investigador quiere estudiar la relación entre dos variables estadísticas. Usando un programa de ordenador obtiene los siguientes valores en un contraste Chi-cuadrado

Pruebas de Chi-cuadrado			
	Valor	gl	Significación (valor p)
Chi-cuadrado de Pearson	157,331(a)	4	,005
N de casos válidos	398		

0 casillas (0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 14,75.

- a. El contraste es estadísticamente significativo, pero el investigador no puede aplicar el contraste Chi-cuadrado porque ninguna celda tiene frecuencia esperada menor que 5
- b. El contraste es estadísticamente significativo, y el investigador puede aplicar el contraste Chi-cuadrado porque se cumplen las condiciones**
- c. El contraste es estadísticamente significativo, pero el investigador no puede aplicar el contraste Chi-cuadrado porque la frecuencia mínima esperada es menor de 20.

Tabla 5.5.2.5. Frecuencias (y porcentajes) de respuestas en el ítem C5 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	71 (76,3)	17 (18,3)	5 (5,4)
b (Verdadera)	47 (50,5)	41 (44,1)	5 (5,4)
c (Falsa)	85 (91,4)	3 (3,2)	5 (5,4)

Los resultados en este ítem sobre la interpretación de los resultados (Batanero y Díaz, 2008) son buenos, siendo sencillo recordar el valor de la frecuencia mínima esperada para aplicar el contraste Chi-cuadrado, así como que se admiten celdas con frecuencia esperada menor que 5 (opciones a y c); fue más difícil reconocer que en este caso se cumplen todas las propiedades de aplicación de este contraste y además el contraste es estadísticamente significativo (opción b), aunque la mitad de la muestra la elige. Posiblemente los estudiantes no recuerdan bien cuáles son estas condiciones o bien no saben si el valor 0.005 es estadísticamente significativo pues no están acostumbrados a trabajar con este nivel de significación.

En esta pregunta un total de 46 estudiantes responden correctamente los tres apartados (50%), 42 estudiantes tienen al menos un error (45%) y 5 estudiantes dejan algún apartado en blanco (5%). Vemos como la mitad saben leer e interpretar la salida de un ordenador y reconocer que se cumplen las condiciones de aplicación del estadístico Chi-cuadrado.

Resultados del ítem 6 del cuestionario 2

- C6. El investigador anterior:
- a. Debe concluir que no hay asociación entre las dos variables en su estudio
 - b. Debe concluir que hay asociación entre las dos variables en su estudio, y la probabilidad de que las variables en realidad sean independientes es menor que 0,005
 - c. Debe concluir que hay asociación entre las dos variables en su estudio; la probabilidad de obtener el valor Chi= 157, 331 si las variables son en realidad independientes es menor que 0,005**

Tabla 5.5.2.6. Frecuencias (y porcentajes) de respuestas en el ítem C6 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	79 (84,9)	10 (10,8)	4 (4,3)
b (Falsa)	56 (60,2)	33 (35,5)	4 (4,3)
c (Verdadera)	32 (34,4)	55 (59,1)	6 (6,5)

En este ítem se evalúa la competencia de los estudiantes para interpretar los resultados del ejercicio anterior y la mayoría fue capaz de reconocer que los resultados llevan a rechazar la hipótesis de asociación entre las dos variables en su estudio (opción a); en consecuencia además de interpretar los resultados, muestran conocimiento de la lógica del contraste de hipótesis.

Los enunciados b) y c) han sido para ellos más complicados, debido a que se añade la interpretación del nivel de significación (se da una interpretación incorrecta en el enunciado b) y la correcta en el c). Recordemos que la interpretación del valor p y nivel de significación causa mucha confusión a estudiantes e investigadores (Vallecillos, 1994; Batanero y Díaz, 2008).

Aunque más de la mitad de la muestra toma como falsa la opción b), son bastante menos los que aceptan la definición correcta del nivel de significación, esto es, la probabilidad de obtener el valor $\chi^2=157,331$ si las variables son en realidad independientes es menor que 0,005 (opción c). Batanero (2000) indica que muchos estudiantes cambian la condición y condicionado en esta definición, porque hay un deseo implícito de poder calcular la probabilidad a posteriori del error. Por ello dan al nivel de significación una interpretación bayesiana y no frecuencial; esta podría ser la causa del error del estudiante en este punto.

En esta pregunta un total de 23 estudiantes responden correctamente los tres apartados (24,7%), 64 tienen al menos un apartado mal (68,8%) y 6 dejan algún apartado en blanco (6,5%).

Resultados del ítem 7 del cuestionario 2

- D1. Si el coeficiente Phi tiene un valor 0,3, esto indica
- Relación inversa y moderada entre las variables
 - Relación directa y moderada entre las variables**
 - Relación directa y alta entre las variables
 - Relación inversa y alta entre las variables

Tabla 5.5.2.7. Frecuencias (y porcentajes) de respuestas en el ítem D1 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	84 (90,3)	9 (9,7)	0 (0)
b (Verdadera)	85 (91,4)	8 (8,6)	0 (0)
c (Falsa)	93 (100)	0 (0)	0 (0)
d (Falsa)	92 (98,9)	1 (1,1)	0 (0)

El ítem sobre la interpretación de un valor Phi fue muy fácil y apenas hubo errores, siendo muy frecuente interpretar que el valor dado no indica una relación alta entre las variables (Ato y López, 1996) (opciones c y d). Un total de 82 estudiantes responden correctamente los tres apartados (88,2%) y 11 tienen al menos una respuesta equivocada (11,8%). Esto nos informa de que los estudiantes han entendido el coeficiente Phi, con su rango y significado, en un alto porcentaje (87,23%).

Resultados del ítem 8 del cuestionario 2

D2. Las medidas de la asociación para tablas 2x2 son:

- a. Coeficiente Phi de Pearson, riesgo relativo y coeficiente de contingencia de Pearson.
- b. Coeficiente Phi de Pearson, riesgo relativo y razón de productos cruzados.**
- c. Riesgo relativo, razón de productos cruzados y V de Cramer.

Tabla 5.5.2.8. Frecuencias (y porcentajes) de respuestas en el ítem D2 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	80 (86)	13 (14)	0 (0)
b (Verdadera)	71 (76,3)	22 (21,2)	0 (0)
c (Falsa)	83 (89,2)	10 (10,8)	0 (0)

En este ítem, también fue muy sencillo reconocer las medidas de asociación para tablas 2x2, rechazando el V de Cramer (opción c) y el coeficiente de contingencia de Pearson (opción a); y recordar que las medidas de la asociación para tablas 2x2 son: coeficiente Phi de Pearson, riesgo relativo y razón de productos cruzados (Ato y López, 1996) (opciones a y b). Un total de 66 estudiantes responden correctamente los tres apartados (71%) y 27 tienen al menos uno mal (29%). Se puede ver como algo más de dos tercios de los estudiantes discriminan las medidas de asociación para tablas 2x2 y tablas rxc.

Resultados del ítem 9 del cuestionario 2

D3. El riesgo relativo por filas y columnas:

- a. Difieren y el valor por filas siempre es mayor o igual que el valor por columnas.
- b. Son diferentes su forma de cálculo, pero coinciden.
- c. Difieren y no hay una regla para saber cual es mayor.

Tabla 5.5.2.9. Frecuencias (y porcentajes) de respuestas en el ítem D3 (n=93)

Respuesta	Correcta	Incorrecta	No responde
A (Falsa)	78 (83,9)	14 (15,1)	1 (1,1)
B (Falsa)	73 (78,5)	19 (20,4)	1 (1,1)
C (Falsa)	61 (65,6)	31 (33,3)	1 (1,1)

Esta pregunta trata de evaluar el conocimiento referido al riesgo relativo en la cual se reconoce con cierta facilidad que las dos primeras opciones son falsas (Ato y López, 1996), pero hay un número alto de errores en la opción (c), ya que los estudiantes recuerdan que por lo general estos valores son diferentes, sin acordarse de que en ocasiones pueden coincidir. Sin embargo, son sólo 19 estudiantes los que responden correctamente los tres apartados (20,4%), 73 tienen al menos un fallo (78,5%) y uno deja algún apartado en blanco (1,1%).

Resultados del ítem 10 del cuestionario 2

D4. La razón de productos cruzados varía de:

- a. $0 < RC < 1$.
- b. $-1 < RC < 1$.
- c. **Puede ser mayor o menor que 1**

Tabla 5.5.2.10. Frecuencias (y porcentajes) de respuestas en el ítem D4 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	63 (67,7)	29 (31,2)	1 (1,1)
b (Falsa)	59 (63,4)	33 (35,5)	1 (1,1)
c (Verdadera)	34 (36,6)	58 (62,4)	1 (1,1)

En esta pregunta, donde se trata de evaluar los conocimientos sobre la variación de la razón de productos cruzados, fue simple recordar que puede tomar valores menores y superiores que 1, siempre positivos (Ato y López, 1996) (opciones a y b). Fue más difícil reconocer que puede ser mayor o menor que uno (opción c). Un total de 28 estudiantes responden correctamente los tres apartados (30,1%), 64 tienen al menos un error (68,8%) y 1 dejan algún apartado en blanco (1,1%). Esta pregunta muestra como los estudiantes tienden a pensar que la medida de asociación tiene que ser menor que 1, regla que no se sigue en este caso.

Resultados del ítem 11 del cuestionario 2

D5. Un investigador quiere relacionar el número de cilindros de una serie de marca de automóviles con el sitio donde se han fabricado. Obtiene las siguientes medidas de asociación en sus datos:

		Valor
Lambda	Número de cilindros dependiente	,183
Lambda	País de origen dependiente	,000
	V de Cramer	,445
	C Coeficiente de contingencia	,532

- Los valores de Lambda indican que se obtiene mejor predicción al predecir el número de cilindros sabiendo el país de origen que al predecir el país de origen sabiendo el número de cilindros.
- Los valores de Lambda indican que se obtiene mejor predicción al predecir el país de origen sabiendo el número de cilindros que al predecir el número de cilindros sabiendo el país de origen.
- Se obtiene igual predicción en ambos casos

Tabla 5.5.2.11. Frecuencias (y porcentajes) de respuestas en el ítem D5 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Verdadera)	36 (38,7)	54 (58,1)	3 (3,2)
b (Falsa)	58 (62,4)	32 (34,4)	3 (3,2)
c (Falsa)	82 (88,2)	8 (8,6)	3 (3,2)

En este ítem sobre la interpretación de valores de Lambda, la opción verdadera es la que menos resultados correctos ha dado; posiblemente los estudiantes no recuerdan el significado de variable dependiente e independiente en este contexto, ya que al estudiar la asociación siempre suponemos que las dos son independientes entre si o asociadas (dependientes).

Lo más sencillo fue recordar que no se obtiene igual predicción en ambos casos (Ruiz-Maya, Martín, Montero y Uriz, 1995), (opción c); es posible que los estudiantes hayan dado esta respuesta recordando esta propiedad, más que leyendo los resultados, pero fue difícil dar la respuesta correcta a las otras dos opciones, es decir interpretar los resultados dados (opciones a y b). Un total de 36 estudiantes responden correctamente los tres apartados (38,7%), 54 tienen al menos un apartado incorrecto (58,1%) y 3 dejan alguno en blanco (3,2%).

Resultados del ítem 12 del cuestionario 2

D6. En el ejemplo anterior, los valores de los coeficientes de asociación V y C indican:

- Una asociación positiva débil
- Una asociación positiva moderada
- Una asociación negativa moderada
- Tendríamos que realizar el contraste Chi-cuadrado para conocer la intensidad y el signo

Tabla 5.5.2.12. Frecuencias (y porcentajes) de respuestas en el ítem D6 (n=93)

Respuesta	Correcta	Incorrecta	No responde
a (Falsa)	86 (92,5)	4 (4,3)	3 (3,2)
b (Falsa)	28 (30,1)	62 (66,7)	3 (3,2)
c (Falsa)	90 (96,8)	0 (0)	3 (3,2)
d (Falsa)	71 (76,3)	19 (20,4)	3 (3,2)

En este ítem sobre la interpretación de los valores V y C , donde todas las opciones son falsas (Ruiz-Maya, Martín, Montero y Uriz, 1995), en general se dan respuestas correctas excepto en el apartado (b). Es decir, se olvida que en las tablas rxc el coeficiente no tiene signo, posiblemente por estar acostumbrados a otros coeficientes con signos, como el de correlación o los estudiados para tablas 2x2. También la mayoría comprende que no es necesario realizar el contraste Chi-cuadrado para conocer la intensidad y el signo (opción d).

Tan sólo 7 estudiantes responden correctamente los tres apartados (7,5%), 83 tienen al menos un error (89,2%) y 3 dejan algún apartado en blanco (3,2%).

5.5.3. SINTESIS DE RESULTADOS EN LOS CUESTIONARIOS

Una vez analizados cada uno de los ítems, se trató de proporcionar una síntesis global, que diese una idea del aprendizaje logrado por los estudiantes. Para ello se realizó un estudio comparado de la dificultad relativa y discriminación de cada subítem una comparación global de los ítems y se analiza el número de preguntas por estudiante en cada cuestionario. A continuación presentamos estos resultados.

Dificultad relativa de las preguntas

El cuestionario elaborado para evaluar el aprendizaje de los estudiantes en la experiencia de enseñanza descrita en el Capítulo 4 podría ser útil como instrumento de evaluación para futuras experiencias. Pero, para ello, conviene asegurar ciertas características psicométricas (Osterlind, 1989; Thorndike, 1989; Martínez, 1995; Martínez, Hernández, Hernández, 2006), una de las cuales se refiere a la inclusión de preguntas con una gama variada de dificultad. Para asegurar este punto, presentamos en la Tabla 5.5.3.1 la dificultad relativa de cada una de las preguntas del cuestionario y en la Figura 5.5.3.1 el histograma correspondiente.

Observamos que los índices de dificultad de la prueba varían desde 0,18 a 0,99, es decir, incluyen ítems muy sencillos y otros bastante difíciles (lo que permitirá graduar la evaluación del aprendizaje). Pero por otro lado, son mayoría los ítems de dificultad moderada o baja, lo que indica unos muy buenos resultados de aprendizaje.

Tabla 5.5.3.1. Índice de dificultad y desviación típica de todos los subítems del cuestionario

	Media	Desviación típica		Media	Desviación típica
A1a	,76	,431	C1c	,92	,268
A1b	,44	,499	C1d	,93	,250
A1c	,54	,501	C2a	,97	,180
A2a	,56	,499	C2b	,36	,483
A2b	,64	,483	C2c	,93	,250
A2c	,41	,494	C3a	,41	,494
A3a	,80	,401	C3b	,87	,340
A3b	,78	,416	C3c	,81	,392
A3c	,70	,459	C2d	,59	,494
A4a	,18	,383	C4a	,53	,502
A4b	,54	,501	C4b	,77	,424
A4c	,37	,486	C4c	,90	,300
A5a	,88	,328	C4d	,78	,416
A5b	,69	,464	C5a	,77	,424
A5c	,64	,483	C5b	,51	,503
A6a	,56	,499	C5c	,91	,285
A6b	,56	,499	C6a	,85	,363
A6c	,18	,383	C6b	,60	,492
B1a	,73	,449	C6c	,33	,473
B1b	,75	,437	D1a	,90	,300
B1c	,30	,483	D1b	,91	,285
B2a	,79	,409	D1d	,99	,105
B2b	,86	,352	D2a	,86	,352
B2c	,65	,480	D2b	,76	,431
B2d	,92	,268	D2c	,89	,314
B3a	,42	,496	D3a	,84	,373
B3b	,19	,392	D3b	,78	,416
B3c	,33	,473	D3c	,67	,473
B4a	,98	,147	D4a	,67	,473
B4b	,97	,180	D4b	,64	,483
B4c	,98	,147	D4c	,36	,483
B5a	,95	,229	D5a	,38	,489
B5b	,86	,352	D5b	,62	,489
B5c	,84	,373	D5c	,89	,314
B6a	,82	,383	D6a	,92	,268
B6b	,47	,502	D6b	,31	,464
B6c	,65	,480	D6c	,97	,180
			D6d	,76	,431

Figura 5.5.3.1. Índices de dificultad en la prueba

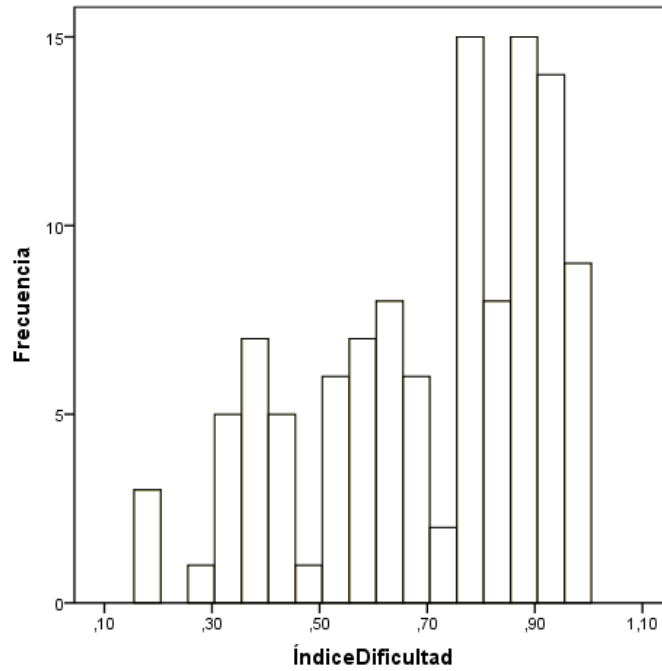


Figura 5.5.3.2. Índices de dificultad en la prueba

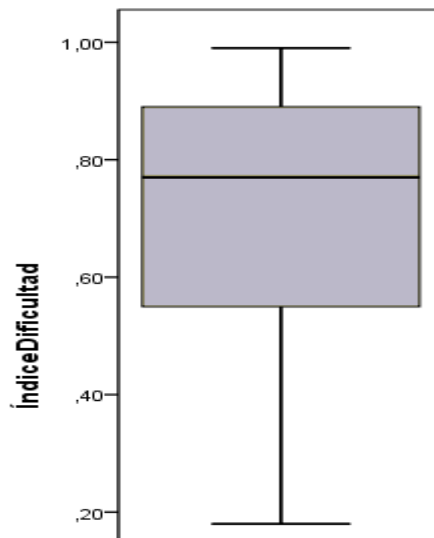


Tabla 5.5.3.2. Estadísticos de índices de dificultad

Estadístico
Media = 0,6865
Mediana = 0,76
Varianza = 0,05
Desviación típica = 0,22423
Mínimo = 0,18
Máximo = 0,99
Rango = 0,81
Asimetría = -0,596
Curtosis = -0,693

Los estadísticos de la distribución de los índices de dificultad se presentan en la Tabla 5.5.3.2, donde vemos una dificultad media de 0,68 y mediana 0,76. En la Figura 5.5.3.2 vemos que la mediana se sitúa alrededor de 0,75 y el primer cuartil por encima del 0,5. Es decir, la mitad de las preguntas es respondida correctamente por tres cuartas partes o más de los estudiantes, y otra cuarta parte del cuestionario es respondida al menos por la mitad de la muestra. La varianza y desviación típica son pequeñas, en relación a la media, lo que indica homogeneidad de resultados.

Comparación global de ítems

También mostramos a modo de resumen la Tabla 5.5.3.3, donde podemos comparar la dificultad relativa de los ítems de los cuestionarios, considerados globalmente, es decir, teniendo en cuenta que todas las preguntas del ítem fuesen correctas; con este criterio han sido clasificados en orden de dificultad. Fueron de dificultad baja para los estudiantes (más de la mitad de la muestra responde correctamente todos los apartados) las preguntas relacionadas con:

Tabla 5.5.3.3. Frecuencias (y porcentajes) de resultados de los cuestionarios

Dificultad para Los estudiantes	Ítem	Todos los apartados correctos	Algún error	Incompleta
Baja	B4	89 (94,7)	3 (3,2)	2 (2,1)
	C1	86 (92,5)	7 (7,5)	0 (0)
	D1	82 (88,2)	11 (11,8)	0 (0)
	B5	72 (76,6)	21 (22,3)	1 (1,1)
	D2	66 (71)	27 (29)	0 (0)
	A3	57 (60,63)	34 (36,17)	3 (3,2)
	A5	56 (59,6)	30 (31,9)	8 (8,5)
	B2	55 (58,5)	37 (39,4)	2 (2,1)
	C4	49 (52,7)	43 (46,2)	1 (1,1)
	Media	C5	46 (49,5)	42 (45,1)
D5		36 (38,7)	54 (58,1)	3 (3,2)
B6		35 (37,2)	54 (57,5)	5 (5,3)
C3		33 (35,5)	60 (64,5)	0 (0)
D4		28 (30,1)	64 (68,8)	1 (1,1)
C2		28 (30,1)	65 (69,9)	0 (0)
A1		28 (29,8)	56 (59,6)	10 (10,6)
Alta	C6	23 (24,7)	64 (68,8)	6 (6,5)
	D3	19 (20,4)	73 (78,5)	1 (1,1)
	A2	15 (16)	79 (84)	0 (0)
	B1	8 (8,5)	83 (88,3)	3 (3,2)
	D6	7 (7,5)	83 (89,3)	3 (3,2)
	A4	6 (6,4)	84 (89,4)	4 (4,3)
	B3	6 (6,4)	84 (89,4)	4 (4,3)
	A6	5(5,3)	88 (93,6)	1 (1,1)

- Determinar cuáles son las celdas que informan del signo de la dependencia en la tabla 2x2 (B4), donde los estudiantes parecen analizar todas las celdas de la tabla, para estudiar el signo de la asociación, razonando al nivel 4 en la clasificación de Pérez Echeverría, y habiendo superado la concepción local de la asociación (Estepa, 1993), que apareció con frecuencia en el estudio de evaluación inicial.
- El modo de calcular los grados de libertad del estadístico Chi-cuadrado (C1), donde los alumnos tienen en cuenta las restricciones y no se observan errores citados por Alvarado (2007) y Olivo (2008).
- La interpretación de un valor del coeficiente de asociación Phi (D1) y reconocer las medidas de asociación para tablas 2x2 (D2). Este tema es nuevo para ellos y aún así, resultó sencillo.
- La diferencia entre asociación y causalidad (B5), que aparentemente es ahora comprendida, lo que supone haber superado la concepción causal de la asociación (Estepa, 1993), que fue resistente en la experiencia de enseñanza del citado autor.
- La relación entre frecuencias dobles absolutas y relativas (A3) y reconocer la representación gráfica adecuada para las tablas rxc (A5), así como la fórmula de cálculo de frecuencias esperadas (B2); sobre estos puntos no tenemos estudios para comparar nuestros resultados.
- La toma de decisión en un contraste, dado un valor p (C4), que es un punto difícil en otras investigaciones, como la de Vallecillos (1996).

Observamos que algunos de estos conocimientos han resultado problemáticos en la investigación previa; por ejemplo, en el trabajo de Estepa (1993) muchos estudiantes muestran la concepción causal de la asociación; por lo cual confunden asociación y causalidad, que ahora parece superado en los participantes en nuestro estudio. Por el contrario resultaron de dificultad alta los siguientes puntos:

- Comprender cuál es el objetivo de construir una tabla de contingencia (A6), posiblemente no se hizo suficiente énfasis en este punto.
- Aceptar la posible igualdad de frecuencias esperadas en caso de asociación, pues algunos estudiantes piensan que nunca pueden ser iguales en este caso (B3). Pensamos que los estudiantes han confundido una condición suficiente (que basta con que alguna de las frecuencias sea diferente para que se dé la independencia) con otra necesaria (que todas han de ser diferentes).

- Recordar o reconocer la fórmula de cálculo de frecuencias relativas condicionales, donde algunos estudiantes no se dan cuenta que no se requiere el total de la muestra, que puede calcularse con las frecuencias dobles (A4). Los estudiantes se limitan a repetir la fórmula aprendida, pero no son capaces de identificar variantes de la misma.
- La interpretación de los coeficientes de asociación V y C (D6) que se asimilan al coeficiente de correlación, que les resulta más familiar, por lo que generalizan indebidamente algunas de sus propiedades.
- Identificar como propiedad de la independencia la igualdad de frecuencias relativas condicionales entre sí o con las marginales, donde el fallo aparece porque el estudiante se contenta con marcar una o dos opciones, siendo las tres verdaderas (B1).
- Suma de las frecuencias relativas marginales, donde algunos estudiantes no se dan cuenta de que es igual a la suma de frecuencias marginales absolutas divididas por el total de la muestra (A2).
- Relación entre el riesgo relativo por filas y el riesgo relativo por columnas (D3), debido al uso específico que se hace en este coeficiente del término variable dependiente e independiente que es distinto del empleado en el resto de la enseñanza.
- Interpretar la definición verbal del valor p , dificultad encontrada en investigaciones previas como la de Vallecillos (1994) (C6).

Han tenido dificultad media los siguientes ítems, donde entre el 25 y el 50% responde correctamente todas las opciones:

- Interpretar si un resultado de un contraste es significativo y saber si se cumplen las condiciones de aplicación del estadístico Chi-cuadrado (C5); posiblemente la dificultad se origina por pedir dos conocimientos diferenciados, cada uno de ellos complejo y además, porque se trabajó con un nivel de significación diferente del utilizado usualmente en las clases:
- Interpretar el valor de Lambda (D5), donde los errores se producen por la confusión ya descrita sobre el significado de variable dependiente e independiente; y diferenciación entre dependencia funcional y aleatoria (B6), donde algunos estudiantes no comprendieron que al variar una variable puede no variar la otra.

- Identificar qué valor p indica mayor asociación en la interpretación de un contraste Chi-cuadrado (C3), debido a que algunos estudiantes confunden significación con asociación.
- Interpretar el rango de la razón de productos cruzados (D4), ya que los estudiantes están acostumbrados a la idea de que un coeficiente de asociación varía entre -1 y +1 o al menos entre 0 y 1.
- Identificar el propósito de los grados de libertad (C2); mientras que los estudiantes entienden que dichos grados de libertad permiten leer la tabla de la distribución y obtener un valor crítico de Chi-cuadrado, no lo identifican como parámetro de dicha distribución, posiblemente porque este concepto es complejo.
- La diferencia entre la información representada en el diagrama de barras adosado y apilado, donde algunos estudiantes no alcanzan un suficiente nivel de lectura de los gráficos, según los niveles definidos por Curcio (1989) (A1).

Discriminación de las preguntas

Para completar la valoración de las características psicométricas del cuestionario, se han calculado también los índices de discriminación, que brinda información acerca de la capacidad para diferenciar entre estudiantes con poco o mucho conocimiento (Barbero, 2003). Si se tratase de una prueba de conocimientos de alguna asignatura, un ítem tiene un alto poder de discriminación si lo aciertan casi todos los que han obtenido alta puntuación en la prueba, y fallan la mayoría de los que han obtenido baja puntuación. Para calcularlo, se seleccionan dos grupos, en función de sus puntuaciones totales de acuerdo con el detalle que a continuación se da (Barbero, 2003):

- Grupo inferior: Se trata del grupo de estudiantes con puntuación inferior al percentil del 30%, es decir, aquellos cuyas puntuaciones son inferiores o iguales al valor 51, que corresponde a dicho percentil.
- Grupo superior: Son los estudiantes cuya puntuación es superior al percentil del 70%, es decir, cuyas puntuaciones son superiores o iguales al valor 57, que corresponde a dicho percentil.

Se considera que el ítem “discrimina”, si existe diferencia significativa entre la proporción de respuestas en ambos grupos. En las Tablas 5.5.3.4 hasta 5.5.3.7 se presenta la diferencia entre estas proporciones, su error típico, el intervalo de confianza de la diferencia y el contraste t de diferencia de proporciones para muestras

independientes (Barbero, 2003). Se observa que la mayoría de los ítems discriminan, siendo excepciones los siguientes para el tema uno, donde hay pocas diferencias en los dos grupos:

- Ítem A1, *b* y *c* (diferencia entre los convenios de representación para el diagrama de barras adosado y apilado). Fue un ítem donde la proporción de respuestas correctas e incorrectas fue muy similar.
- Ítem A2*b* (relación entre frecuencias relativas marginales y total de la muestra); fue muy sencillo en ambos grupos.
- Ítem A4*a* (cálculo de frecuencias relativas condicionales a partir únicamente de las frecuencias dobles), fue excesivamente difícil en los dos grupos.

Tabla 5.5.3.4. Prueba T para la igualdad de medias para los ítems del tema 1

	Apartado	t	g.l.	Sig.	Diferencia de medias	Error típ. De la diferencia	95% I.C. para la diferencia	
							Superior	Inferior
Ítem A1	a	2,03	59	0,05	0,20	0,10	0,00	0,40
	b	0,63	59	0,53	0,08	0,13	-0,18	0,34
	c	-0,67	59	0,51	-0,09	0,13	-0,34	0,17
Ítem A2	a	2,25	59	0,03	0,28	0,12	0,03	0,52
	b	0,36	59	0,72	0,04	0,12	-0,20	0,29
	c	1,19	59	0,24	0,15	0,13	-0,13	0,40
Ítem A3	a	2,03	59	0,05	0,20	0,10	0,00	0,40
	b	2,49	59	0,02	0,27	0,11	0,05	0,49
	c	3,79	59	0,00	0,40	0,11	0,19	0,62
Ítem A4	a	1,15	59	0,25	0,09	0,08	-0,07	0,26
	b	3,45	59	0,00	0,41	0,12	0,17	0,65
	c	1,47	59	0,15	0,18	0,13	-0,07	0,43
Ítem A5	a	1,57	59	0,12	0,14	0,09	-0,04	0,31
	b	1,51	59	0,14	0,17	0,12	-0,06	0,40
	c	2,29	59	0,03	0,27	0,12	0,03	0,51
Ítem A6	a	,36	59	0,72	0,04	0,12	-0,20	0,29
	b	,36	59	0,72	0,05	0,13	-0,21	0,30
	c	-0,39	59	0,70	-0,04	0,10	-0,24	0,16

Para el tema 2 no se encontró discriminación estadísticamente significativa en los siguientes ítems:

- Ítem B1*c* (relación entre las frecuencias relativas condicionales y frecuencias relativas marginales, en caso de independencia). Fue un ítem excesivamente difícil en los dos grupos.
- Ítem B2*d* (relación entre frecuencias esperadas y frecuencias absolutas marginales y el total de la muestra); Ítem B4*c* (conocer que celdas informan de dependencia directa y que celdas informan de una dependencia inversa), Ítem B5, *a* y *c*

(diferencia entre una relación causal y asociación), fue muy sencillo en ambos grupos. Ítem B6c (diferencia entre una dependencia aleatoria y una funcional). Fueron muy sencillos en ambos grupos evidenciando un buen aprendizaje.

Tabla 5.5.3.5. Prueba T para la igualdad de medias para los ítems del tema 2

Ítem	Apartado	t	g.l.	Sig.	Diferencia de medias	Error típ. de la diferencia	95% I.C. para la diferencia	
							Superior	Inferior
Ítem B1	a	3,66	59	0,00	0,40	0,11	0,18	0,63
	b	3,79	59	0,00	0,40	0,11	0,19	0,62
	c	-0,09	59	0,93	-0,01	0,12	-0,26	0,23
Ítem B2	a	2,53	59	0,01	0,26	0,11	0,06	0,48
	b	2,18	59	0,03	0,20	0,09	0,02	0,39
	c	3,47	59	0,00	0,41	0,12	0,17	0,64
	d	0,44	59	0,66	0,04	0,08	-0,13	0,20
Ítem B3	a	3,55	59	0,00	0,41	0,12	0,18	0,65
	b	2,89	59	0,01	0,29	0,10	0,09	0,49
	c	3,14	59	0,00	0,35	0,11	0,13	0,57
Ítem B4	a	1,46	59	0,15	0,07	0,05	-0,02	0,16
	b	1,83	59	0,07	0,10	0,06	-0,01	0,21
	c	0,02	59	0,98	0,00	0,05	-0,09	0,09
Ítem B5	a	0,61	59	0,54	0,03	0,06	-0,08	0,15
	b	2,10	59	0,04	0,17	0,08	0,01	0,33
	c	1,13	59	0,26	0,10	0,09	-0,08	0,29
Ítem B6	a	2,03	59	0,05	0,20	0,10	0,00	0,40
	b	2,23	59	0,03	0,28	0,13	0,03	0,53
	c	0,64	59	0,53	0,08	0,12	-0,16	0,31

Tabla 5.5.3.6. Prueba T para la igualdad de medias para los ítems del tema 3

Ítem	Apartado	t	g.l.	Sig.	Diferencia de medias	Error típ. de la diferencia	95% I.C. para la diferencia	
							Superior	Inferior
Ítem C1	a*				0			
	b*				0			
	c	2,45	59	0,02	0,17	0,07	0,03	0,30
	d	2,45	59	0,02	0,17	0,07	0,03	0,30
Ítem C2	a	-0,98	59	0,33	-0,03	0,03	-0,1	0,03
	b	0,68	59	0,5	0,09	0,13	-0,17	0,34
	c	1,06	59	0,29	0,07	0,06	-0,06	0,2
Ítem C3	a	2,62	59	0,01	0,32	0,12	0,07	0,56
	b	2,18	59	0,03	0,20	0,09	0,02	0,39
	c	2,18	59	0,03	0,20	0,09	0,02	0,39
	d	0,36	59	0,72	0,05	0,13	-0,21	0,30
Ítem C4	a	4,18	59	0,00	0,47	0,11	0,25	0,7
	b	2,76	59	0,01	0,27	0,1	0,07	0,46
	c	1,78	59	0,08	0,13	0,08	-0,02	0,29
	d	1,64	59	0,11	0,17	0,10	-0,04	0,38
Ítem C5	a	1,20	59	0,24	0,14	0,12	-0,09	0,38
	b	1,95	59	0,06	0,25	0,13	-0,01	0,5
	c	2,74	59	0,01	0,20	0,07	0,05	0,35
Ítem C6	a	2,4	59	0,02	0,20	0,08	0,03	0,37
	b	0,00	58	1	0,00	0,13	-0,26	0,26
	c	3,14	59	0,00	0,35	0,11	0,13	0,57

* No se puede calcular T porque las desviaciones típicas de ambos grupos son 0

Para el tema 3 no discriminan los siguientes:

- Ítem C1a y b (conocer la fórmula de los grados de libertad); resultado excesivamente sencilla, ya que todos los estudiantes aciertan.
- Ítem C2 a, b y c (utilidad de los grados de libertad); fue muy sencillo en ambos grupos los apartados a y c (que no sirven para contar las filas y columnas, y que permiten encontrar valores críticos en las tablas de la distribución un valor de Chi-cuadrado, en cambio, resultado más difícil el apartado b (que son un parámetro de la distribución del estadístico).
- Ítem C3d, (conocer que el valor p que indica de forma más concluyente que hay asociación entre las variables es el menor), fue excesivamente difícil en los dos grupos.
- Ítem C6b, interpretar y relacionar el estadístico y el valor p , con una dificultad intermedia en los dos grupos, ya que hay el mismo número de aciertos en ambos grupos.

Tabla 5.5.3.7. Prueba T para la igualdad de medias para los ítems del tema 4

	Apartado	t	g.l.	Sig.	Diferencia de medias	Error típ. de la diferencia	95% I.C. para la diferencia	
							Superior	Inferior
Ítem D1	a	2,45	59	0,02	0,17	0,07	0,03	0,30
	b	2,45	59	0,02	0,17	0,07	0,03	0,30
	c*				0			
	d	1,02	59	0,31	0,03	0,03	-0,03	0,1
Ítem D2	a	2,4	59	0,02	0,20	0,08	0,03	0,37
	b	4,17	59	0,00	0,40	0,1	0,21	0,59
	c	3,02	59	0,00	0,23	0,08	0,08	0,39
Ítem D3	a	1,13	59	0,26	0,10	0,09	-0,08	0,29
	b	-0,27	59	0,79	-0,03	0,10	-0,23	0,17
	c	-0,68	59	0,5	-0,09	0,13	-0,34	0,17
Ítem D4	a	2,21	59	0,03	0,24	0,11	0,02	0,45
	b	2,62	59	0,01	0,31	0,12	0,07	0,54
	c	7,1	59	0,00	0,67	0,1	0,48	0,86
Ítem D5	a	3,17	59	0,00	0,37	0,12	0,14	0,6
	b	1,05	59	0,30	0,13	0,13	-0,12	0,39
	c	2,05	59	0,05	0,17	0,08	0,00	0,33
Ítem D6	a	0,03	59	0,97	0,00	0,06	-0,13	0,13
	b	-0,08	59	0,94	-0,01	0,12	-0,25	0,23
	c	1,46	59	0,15	0,07	0,05	-0,02	0,16
	d	1,57	59	0,12	0,17	0,11	-0,05	0,39

* No se puede calcular T porque las desviaciones típicas de ambos grupos son 0

Para el tema 4 no discriminaron los ítems siguientes:

- Ítem D1 c y d (interpretar el valor $\Phi = 0,3$); Ítem 3 a, b y c (relación entre el riesgo relativo por filas y por columnas); fue muy sencillo en ambos grupos.

- Ítem D5 *b* (interpretación de un valor Lambda). Fue un ítem donde la proporción de respuestas correctas e incorrectas fue muy similar.
- Ítem D6 *a* y *b* (interpretación de los valores *V* y *C*), fue muy fácil en los dos grupos el apartado *a*, y fue excesivamente difícil en los dos grupos el apartado *b*.

Número de respuestas correctas por tema

Para obtener otra indicación global del aprendizaje, se analiza el número total de respuestas correctas, agrupando las preguntas de cada tema, y considerando cada opción de verdadero/falso como un ejercicio, y puntuando 1 punto por cada respuesta correcta. En total obtendríamos para el tema 1, 18 ejercicios; para el tema 2, 19 ejercicios; para el tema 3, 21 ejercicios y para el tema 4, hay 20 ejercicios. En la Tabla 5.5.3.8, mostramos las medias, desviaciones y la diferencia de las puntuaciones obtenidas por este sistema en cada lección.

Observamos que la media obtenida siempre está por encima del valor teórico, lo que confirma de nuevo los buenos resultados de aprendizaje, especialmente en los dos últimos temas (contraste Chi-cuadrado y medidas de asociación) en que la media es bastante más alta que la esperada, a pesar de que requieren mayor nivel de formalización y del gran número de errores descritos en el aprendizaje del contraste de hipótesis (por ejemplo, en Vallecillos, 1994) y Batanero (2000):

Tabla 5.5.3.8. Medias y desviaciones típicas de las puntuaciones en cada tema

	Medias	Desviaciones	Media teórica
Cuestiones tema 1	10,2	2,4	9
Cuestiones tema 2	13,3	2,5	9,5
Cuestiones tema 3	15,8	2,4	10,5
Cuestiones tema 4	15,1	2,4	10

Mostramos en la Figura 5.5.3.3, los gráficos de cajas de las cuatro variables mencionadas. Se observa de nuevo en el gráfico que todos los resultados son buenos, en especial en los temas tres y cuatro, pues las medianas, e incluso los percentiles del 25% se encuentran siempre por encima de los valores esperados de las medias. Ello sugiere que las tres cuartas partes de los alumnos obtienen puntuaciones por encima del valor medio teórico esperado en cada uno de los cuatro temas. Sin embargo, hay valores atípicos que corresponden a estudiantes (pocos casos) con muy malos resultados, lo cual

es razonable, ya que se trató de un grupo- clase completo y la actitud y rendimiento del conjunto de estudiantes no es uniforme.

Tabla 5.5.3.9. Estadísticos descriptivos de la variable puntuación total

	Mínimo	Máximo	Media	Desv. típ.
Puntuación	38	69	54,49	6,24

En la Figura 5.5.3.4 representamos un histograma de las puntuaciones totales (la suma de las cuatro variables), para los estudiantes que completaron ambos cuestionarios (91 estudiantes). En esta variable hemos podido observar (Tabla 5.5.3.9) que se alcanza un valor medio mucho mayor que el valor medio teórico (que sería 35), por lo que los resultados son muy buenos. La variabilidad es pequeña, al comparar el valor de la desviación típica respecto a la media. Todos los estudiantes están sobre la media teórica, es decir, respondieron correctamente al menos a la mitad de los subítems.

Figura 5.5.3.3. Gráficos de cajas de las variables puntuaciones

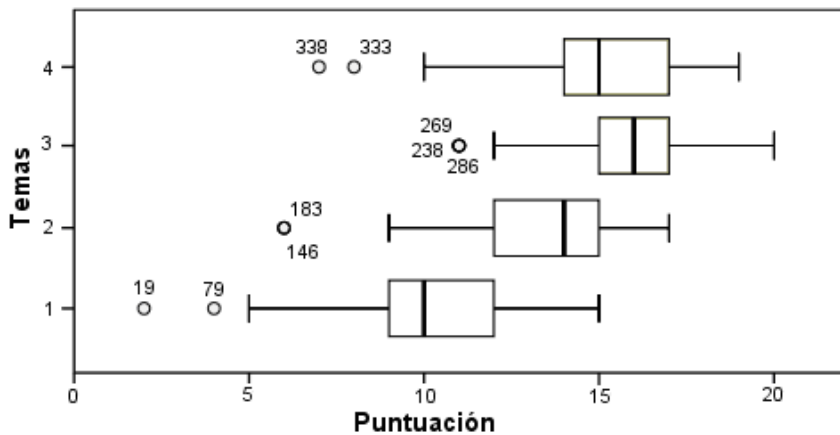
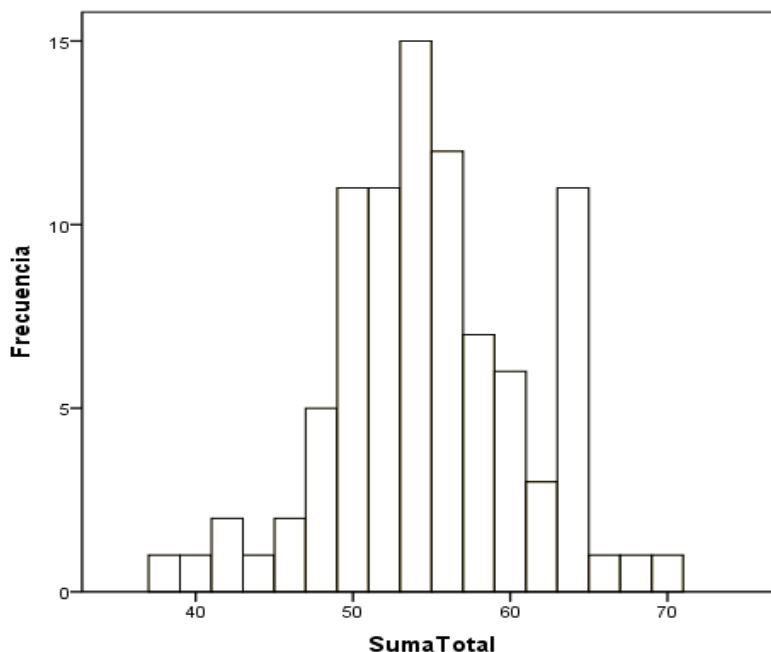


Figura 5.5.3.4. Histogramas de la variable puntuación total



Fiabilidad

Una vez realizado el análisis de ítems, y de las puntuaciones globales, se ha tratado de hacer una estimación de la fiabilidad de la prueba, siendo conscientes que, al ser una primera versión del instrumento, estas evidencias son sólo provisionales y deberán ser completadas en el futuro.

Se ha obtenido el coeficiente de consistencia interna alfa de Cronbach que se basa en el análisis relativo de la varianza total del cuestionario y de las varianzas de los ítems particulares (Barbero, 2003). En el caso de este cuestionario, puesto que la prueba es heterogénea no hay que esperar un índice de consistencia interna muy alto. Se ha calculado este coeficiente a partir de la muestra total de estudiantes ($n = 94$). Para el cálculo del coeficiente alfa se ha utilizado el programa correspondiente del paquete estadístico SPSS. Siendo conscientes de que, como señala Martínez (1995) a mayor tamaño de muestra el coeficiente obtenido será más preciso. No obstante, se cumplen las otras condiciones requeridas por la autora para el cálculo del coeficiente de fiabilidad: representatividad de la muestra, independencia de las respuestas a distintos ítems y administración idéntica al uso pretendido del test. En la Tabla 5.5.3.10 se muestran los resultados obtenidos con el total de ítems del cuestionario.

Tabla 5.5.3.10. Estadísticos de fiabilidad

Alfa de Cronbach	Alfa de Cronbach basada en los elementos tipificados	N de elementos
,705	,722	75

El valor obtenido de Alfa de Cronbach, que es un modelo de consistencia interna, que se basa en la correlación inter-elementos promedio, tiene un valor de 0,705 y el basado en elementos tipificados 0,722. Este valor es razonable, pues supera el valor requerido por diversos autores clásicos, que han dado distintas reglas, desde 0,70, propuesto por Nunnally hasta 0,50 propuesto por Guilford (Santisteban, 1990). Posiblemente, puesto que algunos ítems no discriminan el valor del coeficiente Alfa aumentaría al quitar algunos ítems del cuestionario, lo cual dejamos reflejado para futuras investigaciones.

5.6. RESULTADOS EN LA PRUEBA FINAL

De los 104 estudiantes que participaron en la enseñanza, realizaron el examen final de la asignatura 103 estudiantes (24 hombres y 79 mujeres). Este examen tenía carácter eliminatorio aquellos que no obtuviesen una nota mínima. A continuación se analizan los resultados de los cuatro ítems relacionados con la asociación que se incluyeron en dicho examen, teniendo presente que en este caso los estudiantes saben que sólo hay una opción verdadera en cada ítem, y que en caso de elegir una errónea, tendrán una disminución en la puntuación, lo que puede producir un aumento de no respuestas respecto a las ítems anteriores.

Análisis del ítem 1 en la prueba final

E1. Se quiere estudiar la posible asociación entre si una madre gestante fuma durante su embarazo y si el niño es bajo de peso al nacer. En un estudio se obtienen los siguientes datos:

Gestante	Recién nacido de bajo peso	
	Sí	No
Fumadora	43	207
No fumadora	105	1645

¿Qué conclusiones puede obtener?

- Hay una relación directa entre las variables; se puede observar al comparar las celdas que están en las diagonales sin necesidad de calcular el estadístico Chi-cuadrado.**
- No hay relación pues la mayoría de las madres tiene un niño de peso normal y no fuma durante el embarazo.
- Hay relación inversa pues la mayoría de las madres fumadoras tiene un niño de peso normal y son muy pocas las madres fumadoras con niños bajos de peso.
- No cumple las propiedades de aplicación del estadístico Chi-cuadrado, por lo tanto no podemos dar ninguna conclusión referida a la dependencia.

Tabla 5.6.1. Resultados en el ítem E1

	Número de estudiantes	Porcentaje
No contestan	33	32
a (Verdadera)	35	34
b	20	19,5
c	13	12,6
d	2	1,9
Total	103	100

Como se esperaba la proporción de no respuesta es alta, pues si el alumno se encuentra inseguro prefiere dejarla en blanco a arriesgarse a bajar la nota. El 34% de estudiantes (50% de los que responden) dan la respuesta correcta por lo que reconocen que las celdas de la diagonal principal, es decir, las celdas correspondientes a presencia de los dos caracteres o ausencia de los dos caracteres, informan de una dependencia directa, y que las celdas con presencia de uno de los dos caracteres y ausencia del otro, informan de una dependencia inversa (Inhelder y Piaget, 1955).

Un 20% (28% de los que responde) elige el distractor b), por lo que siguen usando la estrategia de basar su juicio de asociación en una sola celda, esta estrategia es incorrecta, como se describe en muchos estudios psicológicos (Smedlund, 1963; Beyth y Maron, 1982; Shaklee y Mins, 1982; Yates y Curley, 1986). Sería también indicativa de la concepción local de la asociación descrita en el trabajo de Estepa (1993).

Los que eligen el distractor c) son el 12,6% (18,57% de los que responden), usan sólo una distribución de frecuencias, estrategia incorrecta encontrada por Estepa (1993) así como por Smedlund (1963) y Shaklee y Mins (1982). Además, de la baja respuesta al distractor d) se observa como prácticamente la totalidad han aprendido a reconocer los supuestos para la aplicación del estadístico Chi-cuadrado.

Para finalizar el análisis de este ítem, en la Tabla 5.6.2 se presentan las puntuaciones medias en el total de los cuestionarios en los estudiantes que responden o no a cada ítem de la prueba final, con el fin de determinar si hay una relación entre la respuesta al ítem y las respuestas a los cuestionarios

Tabla 5.6.2. Relación entre puntuación al cuestionario y respuesta al ítem 1 de la prueba final

Tema	Puntuación media		D. típica		t	g.l.	Sig. bilateral	Diferencia de medias	Error típ. de la diferencia	95% I.C. para la diferencia	
	Correcta	Incorrecta	Correcta	Incorrecta							
1	10,31	10,10	2,14	2,68	0,40	92	0,69	0,21	0,53	-0,84	1,27
2	13,34	13,31	3,06	2,14	0,07	92	0,94	0,04	0,54	-1,0	1,17
3	16,3	15,52	2,34	2,15	1,64	91	0,11	0,79	0,48	-1,74	0,17
4	15,33	15,02	2,30	2,44	0,62	91	0,54	0,32	0,51	-0,70	1,34

Observamos que los resultados en el ítem 1 no son discriminantes para la puntuación media en ninguno de los cuatro temas. Es decir, no aparece relación estadísticamente significativa entre las puntuaciones medias a los cuestionarios iniciales y la respuesta a este ítem. Una posible conclusión es que, a pesar de la buena preparación formal de los estudiantes, que se observa en las respuestas al cuestionario, al encontrar un ítem como el presentado, donde hay que juzgar la asociación únicamente teniendo en cuenta las celdas de la diagonal, vuelven a sus estrategias intuitivas incorrectas en un porcentaje apreciable a pesar de que el contexto es favorable a las teorías previas.

Análisis del ítem 2 de la prueba final.

E2. Dos variables en una tabla de contingencia son independientes entre sí. Indica cuál de las siguientes propiedades NO SE CUMPLE	
a.	Las frecuencias observadas de todas las celdas son iguales a las frecuencias teóricas esperadas.
b.	En todas las celdas son iguales al producto del total por fila y columna que le corresponda, es decir $h_{i,j} = h_i \cdot h_j$
c.	La suma de las frecuencias de las celdas en la diagonal principal será mayor que la suma de frecuencias en las celdas que no están en esta diagonal.
d.	Las frecuencias relativas condicionales por filas son iguales a las frecuencias relativas marginales por filas.

En este ítem se ha de reconocer diferentes propiedades de la independencia (Batanero y Díaz, 2008). Lo más sencillo para los estudiantes ha sido reconocer la opción verdadera, apartado *c*, reconociendo que en caso de que dos variables sean independientes, no tiene por qué cumplirse que la suma de las frecuencias de las celdas en la diagonal principal será mayor que la suma de frecuencias en las celdas que no están en esta diagonal. De hecho, la comparación de diagonales sería un procedimiento informal de reconocer la dependencia (Inhelder y Piaget, 1955). De los estudiantes que responden, el 73,75% lo hacen correctamente, por lo que han aprendido las propiedades de independencia. La siguiente opción más frecuente es la de los que no responden (22%), siendo pocos los estudiantes que eligen un distractor erróneo.

Tabla 5.6.3. Resultados en el ítem E2

	Número de estudiantes	Porcentaje
No contestan	23	22,3
a	5	4,9
b	9	8,7
c (Verdadera)	59	57,3
d	7	6,8
Total	103	100

Para estudiar si la respuesta a este ítem se relaciona con los resultados en el cuestionario, en la Tabla 5.6.4 presentamos los resultados del ítem 2, que no son discriminantes para la puntuación media en ninguno de los cuatro temas. Es decir, no aparece relación estadísticamente significativa entre las puntuaciones medias a los cuestionarios iniciales y la respuesta a este ítem. En este caso hubo un alto porcentaje de respuestas correctas, de lo que se deduce la falta de discriminación.

Tabla 5.6.4. Relación entre puntuación al cuestionario y respuesta al ítem 2 de la prueba final

Tema	Punt. Media		D. típica		t	gl	Sig. bilateral	Diferencia de medias	Error típ. de la diferencia	95% I.C. para la diferencia	
	Correcta	Incor.	Correcta	Incor.							
1	10,09	10,32	2,39	2,65	0,45	92	0,65	0,24	0,53	-0,81	1,28
2	13,49	13,05	2,43	2,64	-0,83	92	0,41	-0,44	0,53	-1,49	0,62
3	15,85	15,72	2,24	2,26	-0,28	91	0,78	-0,13	0,47	-1,07	0,81
4	15,31	14,87	2,13	2,66	-0,89	91	0,38	-0,44	0,5	-1,43	0,54

Análisis del ítem 3 de la prueba final.

E3. Un psicólogo pone a punto una terapia para disminuir la ansiedad. Recoge datos de 200 pacientes a 100 de los cuáles (elegidos al azar) aplica la terapia y al resto otra alternativa y pasados tres meses de tratamiento analiza la posible asociación entre terapia seguida (nueva o antigua) y disminución del nivel de ansiedad (disminuye o no disminuye). Aplica un contraste Chi-cuadrado de independencia y obtiene un valor $\chi^2=49,1$, con 1 g.l. ¿Qué conclusión debiera obtener de estos resultados?

- El resultado del contraste no es estadísticamente significativo, pues el valor en la tabla así lo confirma al comparar con el obtenido por el investigador. Debiera concluir que las variables no están relacionadas entre sí.
- El resultado del contraste es estadísticamente significativo, pues $p < 0,05$. Debiera concluir que las variables están relacionadas entre sí. La relación es muy pequeña, porque el valor p es muy bajo.
- El resultado del contraste es estadísticamente significativo, pues $p < 0,05$. Debiera concluir que las variables están relacionadas entre sí. Para saber si la relación es fuerte o débil tenemos que calcular alguna medida de asociación.**
- No podemos concluir si hay relación entre las variables sólo con el valor Chi-cuadrado. Tenemos que calcular los coeficientes de asociación.

Tabla 5.6.5. Resultados en el ítem E3

	Número de estudiantes	Porcentaje
No contestan	49	47,6
a	4	3,9
b	4	3,9
c (Verdadera)	39	37,9
d	7	6,7
Total	103	100

En este ítem se pide interpretar los resultados de un contraste Chi-cuadrado de independencia (Batanero y Díaz, 2008). De los estudiantes que responden, el 37,8% lo hacen correctamente, por lo que han aprendido la interpretación del estadístico y la regla

de decisión de un contraste de independencia, además de comprender que se requiere una medida de asociación para evaluar la intensidad de la misma.

No obstante, casi la mitad de los estudiantes deja la respuesta en blanco, lo que muestra su inseguridad al elegir una única respuesta, posiblemente porque en algunos de los distractores parte del enunciado es verdadero. Con ello, el porcentaje de respuestas correctas es inferior a la mitad de los estudiantes. Habría que considerar los que dan las opciones b y d que son parcialmente correctas (un 10% entre las dos), lo que mejoraría los resultados en este ítem. Son muy pocos los que piensan que el resultado no es estadísticamente significativo (opción a).

Para analizar la relación de este ítem con los resultados en el cuestionario, en la Tabla 5.6.6 presentamos los resultados del ítem 3, que no son discriminantes para la puntuación media en los temas 2 y 4, pero sí en el 1.

Tabla 5.6.6. Relación entre puntuación al cuestionario y respuesta al ítem 3 de la prueba final

Tema	Punt. media		D. típica		t	g.l.	Sig. bilateral	Diferencia de medias	Error típ. de la diferencia	95% I.C. para la diferencia	
	Correcta	Incor.	Correcta	Incor.							
1	10,86	9,78	2,14	2,59	-2,07	92	0,04	-1,08	0,52	-2,11	-0,05
2	13,46	13,24	2,41	2,58	-0,41	92	0,68	-0,22	0,54	-1,29	0,85
3	16,18	15,58	2,58	2,01	-1,25	91	0,22	-0,6	0,48	-1,56	0,35
4	15,38	14,98	2,13	2,49	-0,78	91	0,44	-0,4	0,51	-1,41	0,61

Deducimos que el aprendizaje de los diferentes tipos de frecuencia en la tabla (Tema 1) y el cálculo del estadístico Chi-cuadrado y contrastes asociados (independencia y homogeneidad, Tema 3) influyó en la respuesta final del ítem 3, por lo que los estudiantes han transferido su aprendizaje de los citados temas al responder a este ítem. Ello es razonable pues el ítem plantea la interpretación de los resultados de un contraste Chi-cuadrado, relacionado con los dos temas anteriores.

Análisis del ítem 4 de la prueba final.

E4. Un investigador quiere relacionar el número de cilindros de una serie de marca de automóviles con el sitio donde se han fabricado (Europa, Estados Unidos o Japón). Obtiene las siguientes medidas de asociación en sus datos: V de Cramer= 0,445; Coeficiente de contingencia de Pearson= 0,532; Lambda de Goodman y Kruskal (Variable dependiente= Número de cilindros)= 0,183.

¿Qué conclusión debe obtener el investigador?

- Hay una relación muy alta entre las variables y se puede predecir muy fácilmente el número de cilindros del coche, si se sabe donde se ha fabricado
- Hay una relación moderada-alta entre las variables y se puede predecir muy fácilmente el número de cilindros del coche, si se sabe donde se ha fabricado
- Hay una relación moderada-alta entre las variables, pero no es fácil predecir el número de cilindros del coche, si se sabe donde se ha fabricado**
- Hay una relación muy alta entre las variables, pero no es fácil predecir el número de cilindros del coche, si se sabe donde se ha fabricado

Tabla 5.6.7. Resultados en el ítem E4

	Número de estudiantes	Porcentaje
No contestan	42	40,7
a	0	0
b	15	14,6
c (Verdadera)	46	44,7
d	0	0
Total	103	100

En este ítem, en el cual se pide a los estudiantes la interpretación de valores de los coeficientes de asociación V , C y de Lambda (Ruiz-Maya, Martín, Montero y Uriz, 1995), de nuevo la proporción de no respuesta es muy alta.

La opción verdadera, elegida por los estudiantes que han aprendido la interpretación de estos valores es la que tiene mayor frecuencia, pero no alcanza la mitad de la muestra, pero representa el 75'4% de los que han respondido al ítem. Hay que tener en cuenta la opción b, que es parcialmente correcta, respondida por otro 15% de estudiantes, 24'6% de los que responde, lo que mejoraría los resultados. También destacamos que la interpretación de la intensidad los valores V y C , no han tenido ningún error, ya que las opciones a y d , tienen una frecuencia de 0. Los pocos errores (15 sujetos), son los que han confundido la variable dependiente y la independiente al interpretar el valor de Lambda.

Finalmente, en la Tabla 5.6.8 se presentan los resultados en el ítem 4, para analizar la relación de esta pregunta con las respuestas al cuestionario. No aparece relación con ninguno de los cuestionarios.

Tabla 5.6.8. Relación entre puntuación al cuestionario y respuesta al ítem 4 de la prueba final

Tema	Puntuación media		D. típica		t	g.l.	Sig. bilateral	Diferencia de medias	Error típ. de la diferencia	95% I.C. para la diferencia	
	Correcta	Incor.	Correcta	Incor.							
1	10,66	9,81	2,25	2,60	-1,66	92	0,10	-0,85	0,51	-1,86	0,17
2	13,37	13,28	2,37	2,63	-0,16	92	0,88	-0,08	0,52	-1,12	0,96
3	15,76	15,82	2,34	2,18	0,13	91	0,9	0,06	0,47	-0,87	0,99
4	15,07	15,18	2,32	2,41	0,21	91	0,83	0,11	0,49	-0,88	1,00

En resumen, los resultados en la prueba final no indican una relación concluyente con las evaluaciones intermedias, lo que podría deberse a un efecto de olvido, pues las primeras evaluaciones se realizaron al poco tiempo de finalizada la enseñanza. Por otro lado, hay que comprender que en el examen final los estudiantes se examinaron de la materia completa, mientras que en las autoevaluaciones se concentraron en el tema de la asociación. Finalmente la penalización (perder puntos por respuestas erróneas) ha sesgado algo los resultados, pues algunos estudiantes no se

atreveron a responder en caso de duda, efecto que no aparece en las autoevaluaciones. En todo caso destacamos que la mayor influencia se da en el ítem relacionado con la interpretación del contraste de independencia y el estadístico Chi-cuadrado (ítem 3).

5.7. CONCLUSIONES

En este capítulo se presentaron los resultados de la parte de la evaluación que corresponde a las preguntas cerradas, tanto en los cuestionarios pasados durante la enseñanza, como en la prueba final de la asignatura. Nuestros objetivos fueron estimar la proporción de estudiantes en la muestra que comprenden y aplican los objetos matemáticos incluidos en la enseñanza. Para ello se construyó un cuestionario, siguiendo las recomendaciones psicométricas habituales (Osterlind, 1989; Thorndike, 1989; Martínez, 1995; Martínez, Hernández, Hernández, 2006), que contenía ítems de los cuatro temas estudiados. Igualmente se pasaron cuatro ítems de opción múltiple en el examen final de la asignatura. A continuación se hace un resumen de las principales conclusiones.

El análisis a priori del cuestionario indicó que se han tenido en cuenta la mayor parte de los objetos incluidos en la enseñanza, asegurando la validez de contenido al proceso de evaluación. Por otro lado, las características de dificultad y discriminación de ítems, así como valores adecuados de fiabilidad, determinadas a partir de los resultados en la muestra de estudiantes, sugieren una alta idoneidad didáctica del cuestionario, según los criterios explicitados por Batanero y Díaz (2005). Estas autoras interpretan los criterios clásicos de calidad de los instrumentos de evaluación en el siguiente sentido:

- La dificultad de un ítem daría una medida de su idoneidad cognitiva; es decir del grado de representatividad de los significados evaluados respecto a los significados personales. En nuestro estudio, el 75% de los ítems fueron respondidos por el 50% o más de los alumnos, lo que indica una alta idoneidad cognitiva.
- La discriminación de un ítem valoraría su idoneidad evaluadora, un componente de la idoneidad instruccional, según Batanero y Díaz (2005). En nuestro caso, la mayoría es discriminante, pero no todos, de modo que algunos ítems se podrían eliminar en un futuro instrumento o tratar de mejorar su capacidad discriminante
- La validez de contenido de un cuestionario indicaría una idoneidad epistémica, o grado de representatividad del instrumento en cuanto al significado objeto de evaluación. Sería muy alta en nuestro estudio, al contener el cuestionario la mayoría

de los objetos matemáticos considerados en la enseñanza y un número similar de ítems para cada tema.

- La fiabilidad daría una medida de la estabilidad de la respuesta, es decir sería otro componente de la idoneidad evaluadora; en nuestro caso es suficiente, dado que la muestra no es muy amplia, posiblemente pueda aumentar en otras aplicaciones del cuestionario.
- La validez externa sugeriría una idoneidad generalizadora o externa en cuanto los resultados se generalizarían a otros estudiantes. Pensamos que en nuestro caso sólo se puede generalizar a estudiantes semejantes a los de la muestra, es decir, alumnos de psicología con un programa de enseñanza similar.

Los resultados obtenidos sugieren unos buenos resultados de aprendizaje por parte de los estudiantes participantes en la enseñanza. El número total de respuestas correctas alcanza un valor medio mucho mayor que el valor medio teórico (que sería 35), por lo que los resultados de aprendizaje global fueron muy buenos. La variabilidad es pequeña, al comparar el valor de la desviación típica respecto a la media. Todos los estudiantes están sobre la media teórica, es decir, respondieron correctamente al menos a la mitad de los subítems. Este mismo resultado se observa en las puntuaciones parciales a los ítems en cada uno de los cuatro temas, especialmente en los dos últimos temas (contraste Chi-cuadrado y medidas de asociación) en que la media es bastante más alta que la esperada, a pesar de los temas que requieren mayor nivel de formalización.

Este aprendizaje se evidencia también en el análisis detallado de las respuestas en cada ítem, y comparando, cuando es pertinente con la investigación previa. Entre otros resultados del análisis de ítems, resaltamos los siguientes:

- Los estudiantes parecen razonar en el nivel 5 de Pérez Echeverría y en consecuencia, han adquirido estrategias de nivel bastante más alto que las empleadas en el estudio de evaluación inicial para analizar la asociación en tablas de contingencia (ítem B1). Asimismo casi sin excepción han aprendido a diferenciar las celdas que informan de dependencia directa y dependencia inversa (ítem B4).
- La mayoría de los estudiantes parece haber superado la concepción causal de la asociación descrita por Estepa (1993), en cuyo estudio fue resistente con la instrucción (respuestas al ítem B5); también reconocen que en la dependencia

aleatoria a un valor de la variable independiente pueden corresponder varias de la dependiente (B6).

- Los estudiantes han aprendido la fórmula de cálculo de los grados de libertad, a pesar que no se les exigía. Apenas aparece el error citado por Alvarado (2007) y Olivo (2008) consistente en no restar el número de restricciones en el cálculo (C1).
- Fue sencillo reconocer el valor de la frecuencia mínima esperada y la proporción de celdas con baja frecuencia para aplicar el contraste Chi-cuadrado (C5)). Por tanto, los estudiantes parecen reconocer las condiciones en que pueden aplicar este estadístico en una tabla de contingencia.
- La mayoría de estudiantes interpreta correctamente el valor Phi (D1), así como los coeficientes de asociación V y C (D6), habiendo adquirido comprensión suficiente de estos estadísticos, los valores que toman y su interpretación.

Por otro lado, la dificultad de los ítems fue variada, encontrando baja proporción de respuestas correctas en ítems relacionados con los siguientes puntos:

- Algunos estudiantes no alcanzan el nivel de *extracción de datos*, que consiste en poner en relación un elemento de un eje de un gráfico con el de otro eje, en este caso leer la frecuencia asociada a un valor de la variable (Bertin, 1967) (ítem A1)
- Se comprende mal cuál es el objetivo de construir una tabla de contingencia. Hubiese sido necesario insistir más en su interés como resumen estadístico de los datos e instrumento de análisis.
- Algunos estudiantes olvidan las fórmulas de cálculo de frecuencias relativas condicionales, y no se dan cuenta que no se requiere el total de la muestra, pues puede calcularse con las frecuencias dobles. En otros casos no recuerdan que la suma de las frecuencias relativas marginales, es igual a la suma de frecuencias marginales absolutas divididas por el total de la muestra. Este resultado nos ha sorprendido pues pensábamos a priori que no se trataba de temas complejos; posiblemente ello hizo que no se le prestase suficiente atención o bien por no haber realizado los cálculos a mano, sino sólo a través de la tabla Excel.
- Aceptar la posible igualdad de frecuencias esperadas en caso de asociación, pues algunos estudiantes piensan que nunca pueden ser iguales en este caso. La formulación por nuestra parte de esta propiedad hizo que los estudiantes hiciesen una generalización indebida.

- También fue difícil reconocer que la razón de productos cruzados puede tomar valores superiores a la unidad (igualmente, los conocimientos sobre el coeficiente de correlación y otros coeficientes de asociación, cuyo valor absoluto está limitado por la unidad confundieron a los estudiantes en este punto) (D4).
- Relación entre el riesgo relativo por filas y el riesgo relativo por columnas, pues los estudiantes no supieron pasar de uno a otro y no reconocen que a veces pueden coincidir (D3). También hubo dificultad en interpretar el valor Lambda (D5), por falta de discriminación de la variable dependiente e independiente.

Los resultados en la prueba final, sin embargo, fueron peores de lo esperado, posiblemente porque algunos estudiantes temerosos de la penalización que reciben en caso de respuesta incorrecta, dejaron algunos sin responder (lo que se observa en el alto número de respuestas en blanco). Por otra parte, aunque el aprendizaje fue muy bueno a corto plazo, hay un olvido transcurrido unos meses, ya que las respuestas correctas en ningún caso llegan al 50%, siendo mejores los obtenidos en el ítem relacionado con las propiedades de la independencia estadística.

Tampoco parece haber una alta relación entre las respuestas al cuestionario y a los ítems, excepto en lo relacionado con el tema 3 (contraste Chi-cuadrado) que fue más discriminativo y precisamente en el ítem relacionado con la interpretación de este contraste, así como en otro relacionado con un juicio de asociación en una tabla de contingencia. En todo caso, estos resultados en la prueba final hay que interpretarlos con precaución, pues se basan tan sólo en 4 ítems (frente a los 24 de los cuestionarios). Sería necesaria una mayor muestra de ítems para comprobar que el aprendizaje no fue tan efectivo en el largo plazo; pero ello requerirá una nueva investigación.

Por último, destacamos que todos estos resultados de la evaluación del aprendizaje es una aportación de este trabajo de investigación de destacado valor para futuras planificaciones de la enseñanza, así como para la evaluación de los aprendizajes adquiridos.

CAPÍTULO 6

EVALUACIÓN DEL APRENDIZAJE MEDIANTE PROBLEMAS ABIERTOS

- 6.1. Introducción
- 6.2. Objetivos de la evaluación
- 6.3. Análisis a priori del significado evaluado en los ítems abiertos
 - 6.3.1. Análisis del problema A: Contraste de homogeneidad
 - 6.3.2. Análisis del problema B: Medidas de asociación en tablas 2x2
 - 6.3.3. Análisis del problema C: Medidas de asociación en tablas rxc
 - 6.3.4. Análisis del problema D: Cálculo de probabilidades y juicios de asociación
 - 6.3.5. Análisis del problema E: Juicio de asociación y justificación
- 6.4. Síntesis de las evaluaciones mediante problemas abiertos
- 6.5. Metodología de análisis de los ítems abiertos
- 6.6. Resultados en el problema A: Contraste de homogeneidad
 - 6.6.1. Formulación de hipótesis
 - 6.6.2. Estadístico de contraste y probabilidad asociada
 - 6.6.3. Toma de decisión
 - 6.6.4. Interpretación de resultados
- 6.7. Resultados en el problema B: Medidas de asociación en tablas 2x2
 - 6.7.1. Coeficiente Phi de Pearson
 - 6.7.2. Riesgo relativo
 - 6.7.3. Razón de Productos cruzados
- 6.8. Resultados en el problema C: Medidas de asociación en tablas rxc
 - 6.8.1. Coeficiente de Contingencia de Pearson
 - 6.8.2. V de Cramer
 - 6.8.3. Lambda de Goodman y Kruskal
- 6.9. Resultados en el problema D: Cálculo de probabilidades y juicios de asociación
 - 6.9.1. Cálculo de Probabilidad
 - 6.9.2. Juicio de asociación
 - 6.9.3. Estrategia en el juicio de asociación
- 6.10. Resultados en el problema E: Juicio de asociación y justificación
 - 6.10.1. Juicio de asociación
 - 6.10.2. Signo de la asociación
 - 6.10.3. Intensidad de la asociación
 - 6.10.4. Justificación del juicio de asociación
- 6.11. Síntesis de resultados en los problemas abiertos
- 6.12. Relación entre conocimiento teórico y resolución de problemas
- 6.13. Conclusiones

6.1. INTRODUCCIÓN

Como se ha indicado en los capítulos anteriores, entre las pruebas de evaluación pasadas a los estudiantes se incluyeron algunos problemas abiertos, algunos de los cuáles los resolvieron los estudiantes en el aula de informática, con el programa en Excel suministrado para el estudio de las tablas de contingencia y otros sin esta ayuda.

En este capítulo se analizan las respuestas a estos problemas abiertos, para caracterizar de una forma más profunda, al igual que se hizo en la investigación de Tauber (2001) la comprensión lograda por los estudiantes de los diferentes elementos de significados relacionados con la tabla de contingencia, incluidos en el proceso de estudio descrito en el Capítulo 4. En particular, se trata de evaluar su capacidad de argumentación y las estrategias utilizadas al resolver los problemas. Como también se ha indicado, se entiende la evaluación como un estudio de la correspondencia entre el *significado institucional* implementado en la enseñanza y el *significado personal* efectivamente alcanzado por los estudiantes (Godino, 2002b).

El resultado de la evaluación servirá también para valorar en el capítulo de conclusiones la idoneidad didáctica (epistémica, cognitiva, semiótica, emocional y mediacional) del proceso didáctico propuesto para los estudiantes a quienes va dirigida (Godino, Contreras y Font, 2006). En lo que sigue se analizarán los instrumentos de recolección de datos y presentarán los resultados del estudio.

6.2. OBJETIVOS DE LA EVALUACIÓN

El *objetivo principal* al construir las tareas abiertas fue completar la evaluación realizada con el cuestionario y descrita en el Capítulo 5 con indicadores de la capacidad de aplicación de lo aprendido a la resolución de problemas relacionados con las tablas de contingencia. De este objetivo principal se deducen otros objetivos:

- O1: El primero de ellos consiste en estimar la proporción de estudiantes en la muestra que comprenden y particularizan los objetos matemáticos incluidos en la enseñanza en un contexto de aplicación de los mismos, en que, en lugar de elegir de una serie de respuestas pre-establecidas, pueden utilizar sus propias estrategias y argumentos. De este modo podemos evaluar una comprensión más profunda que la llevada a cabo con los ítems de opción múltiple (Capítulo 5).
- O2: Asimismo, deseamos identificar cuáles de ellos resultaron difíciles, identificando los conflictos semióticos (Godino, Batanero y Font, 2007) presentados respecto a los mismos, mostrando tanto las características en la comprensión en el grupo de estudiantes, como su variabilidad.

Todo ello permitirá completar la caracterización del significado personal de los estudiantes sobre las tablas de contingencia iniciado en el Capítulo 5 y compararlo con el pretendido en la enseñanza (Capítulo 3).

6.3. ANÁLISIS A PRIORI DEL SIGNIFICADO EVALUADO EN LOS ÍTEMS ABIERTOS

A continuación se analizan cinco problemas abiertos (Anexo 2), identificando su contenido y las respuestas esperadas de los estudiantes. Los tres primeros fueron resueltos durante las sesiones de prácticas, utilizando el software proporcionado. Además se incluyen otros dos problemas abiertos que fueron resueltos como parte de las evaluaciones parciales, con ayuda tan sólo de la calculadora. Cada uno de ellos se componen de un enunciado inicial y varios apartados donde se realizan preguntas encaminadas a evaluar diferentes conocimientos de los estudiantes.

En el proceso de elaboración de las tareas y al igual que en el cuestionario se han seguido una serie de recomendaciones psicométricas habituales (Osterlind, 1989; Thorndike, 1989; Martínez Arias, 1995):

- En primer lugar, se delimitó el contenido a evaluar con los problemas, a partir del análisis del significado de referencia llevado a cabo en la Sección 1.5 y del de los materiales diseñados para la enseñanza, que se describen en el Capítulo 4.
- Se procedió a la elaboración de una colección de problemas iniciales. A partir de ellos, mediante una serie de revisiones y analizando en cada una los contenidos evaluados, se seleccionaron, posteriormente, los que habrían de constituir los instrumentos.

Los problemas fueron tomados de diferentes libros de texto o investigaciones previas. Una vez que se llegó a aquellos que parecían cubrir el contenido pretendido, se procedió a homogeneizar la redacción y a cambiar el contexto cuando éste no fuese familiar al estudiante. Se hicieron pruebas de comprensión del enunciado, con algunos estudiantes de la misma especialidad y curso que los que habría de responderlos, modificando la redacción en los casos que fue necesario, por no estar suficientemente claro el enunciado. Se tuvieron también en cuenta las recomendaciones para mejorar la redacción: Evitar detalles innecesarios, relevancia de las preguntas formuladas para el estudio, nivel de lectura adecuado, brevedad, claridad y nivel apropiado de abstracción.

También se ha tenido en cuenta la dimensión de la tabla, incluyendo tres problemas referidos a tablas 2x2, uno que trata de una tabla 2x3 y otro sobre una tabla 3x3. Se varía asimismo la intensidad de la asociación, incluyendo un caso de independencia, se tiene en cuenta los diversos tipos de relaciones que pueden explicar la asociación entre variables (Barbancho, 1973) y se combinan casos en que el alumno

tiene o no tiene teorías previas sobre dicha relación.

6.3.1. ANÁLISIS DEL PROBLEMA A: CONTRASTE DE HOMOGENEIDAD

Problema A. Para un experimento sobre memoria a largo plazo un psicólogo selecciona 3 palabras y decide evaluar el componente emocional de las mismas con una muestra aleatoria. Presenta, independientemente, cada palabra a 100 sujetos y registra si la palabra es percibida como emocionalmente positiva (agradable, placentera, etc.), negativa (desagradable, no placentera, etc.) o neutra. A la vista de los resultados obtenidos, ¿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.

(Adaptado de Pardo y San Martín, 1998, pág. 554)

Solución prevista para el apartado (a). Formulación de hipótesis

En este problema, se analiza una tabla 3x3, donde existe una asociación muy pequeña (coeficiente $C=0,187$) que llega a ser estadísticamente significativa ($p=0,029$). El alumno no tiene teorías previas al respecto y en caso de ser las muestras no homogéneas, la relación sería de tipo causal, pues serían las palabras las que influyen en el componente emocional.

Se plantea al estudiante un contraste de homogeneidad, guiándole en la realización del contraste mediante una serie de preguntas. Para resolver el primer apartado el estudiante ha de leer e interpretar el enunciado, recordar que en un contraste de homogeneidad se trata de decidir si los datos provenientes de muestras de diferentes poblaciones, pueden considerarse o no equivalentes (Batanero y Díaz, 2008). El estudiante ha de leer la tabla, identificar la variable del estudio (grado de componente emocional) y las poblaciones en estudio (reacciones de una muestra de sujetos a tres palabras; se trata de tres poblaciones relacionadas o lo que es lo mismo tres mediciones sobre los mismos sujetos). Una posible formulación correcta de las hipótesis sería la siguiente, donde se establece como hipótesis nula la de no diferencia y como alternativa su contraria:

H_0 : Las tres palabras tienen idéntica distribución de componente emocional.

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

Admitiremos cualquier variante lingüística de esta formulación, por ejemplo:

H_0 : Las tres distribuciones 1, 2 y 3 son homogéneas.

H_1 : Alguna distribución es diferente.

Solución prevista para el apartado (b). Estadístico de contraste y probabilidad asociada

En el segundo apartado, el estudiante ha de elegir un estadístico de contraste, siendo el adecuado en este caso el Chi-cuadrado (Lipschutz y Schiller, 1999; Ruiz-Maya y Martín, 2002), especificando sus grados de libertad. Utilizando el programa Excel proporcionado, deberá introducir los datos y calcular la probabilidad de obtener un valor más extremo o igual que el obtenido, en caso de homogeneidad de las poblaciones.

Tabla 6.3.1.1. Componentes de Chi-cuadrado proporcionado por el software

	1	2	3	Total
Positivo	2,023	3,314	0,159	5,495
Negativo	0,003	0,88	0,993	1,876
Neutro	2,253	0,853	0,333	3,44
Total	4,279	5,047	1,485	10,81

El programa calcula los componentes de la Chi-cuadrado y su valor total, proporcionando la Tabla 6.3.1.1. También calcula el valor Chi cuadrado, que toma el

valor $\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 10,81$. El programa no especifica los grados de libertad,

que deben ser proporcionados por el estudiante, recordando la relación g.l. = (número de filas - 1) x (número de columnas - 1) = 2 x 2 = 4. Una vez proporcionados los grados de libertad, el programa proporciona automáticamente el p -valor = 0,029.

Solución prevista para el apartado (c). Toma de decisión

Para resolver este apartado el estudiante podría leer en las tablas de la distribución Chi-cuadrado el valor crítico para $\alpha=0,01$, lo que implica un manejo correcto de las mismas. Este valor es $\chi^2_{4;0,01} = 13,28$. Debe comparar este valor crítico con el valor obtenido de la tabla de datos. Como $\chi^2_{\text{exp}} < \chi^2_{4;0,01}$, el estudiante no podría rechazar la hipótesis nula H_0 , en el contraste de homogeneidad (Batanero y Díaz, 2008).

Otra solución alternativa sería razonar que el p -valor = 0,029 > $\alpha=0,01$ no es

estadísticamente significativo, para el nivel de significación especificado en el problema, aunque lo sea para $\alpha=0,05$. Ello implica recordar la definición de valor p y nivel de significación, así como la relación que han de guardar para rechazar la hipótesis nula H_0 . En cualquiera de los dos casos, el estudiante debe recordar la lógica del contraste de hipótesis y la forma en que se construyen la región de aceptación y rechazo, bien a partir del valor crítico o del p -valor.

Solución prevista para el apartado (d). Interpretación de resultados

El último paso es la interpretación de los resultados en el contexto del problema. Como hemos visto, el valor p obtenido en el contraste es mayor que el nivel de significación; esto quiere decir que no podemos rechazar la hipótesis nula con un nivel de significación del 0,01. En el supuesto planteado, implica que a nivel 0,01 no podemos suponer que alguna de las palabras tenga diferente componente emocional y en consecuencia, a dicho nivel podríamos provisionalmente aceptar la homogeneidad de las poblaciones (Batanero y Díaz, 2008).

6.3.2. ANÁLISIS DEL PROBLEMA B: MEDIDAS DE ASOCIACIÓN EN TABLAS 2X2

Problema B. Se desea estudiar hasta qué punto existe relación entre el tiempo de residencia de inmigrantes en nuestro país y su grado de ansiedad. Se dispone de una muestra de 207 inmigrantes a los que se les evaluó en ambas variables obteniéndose la siguiente tabla de frecuencias observadas.

Tiempo de Residencia	Grado de ansiedad		Total
	Bajo	Alto	
Menos tiempo	36	81	117
Más tiempo	81	9	90
Total	117	90	207

Calcule e interprete las medidas de asociación para tablas 2x2
(Adaptado de Guàrdia, Freixa, Però, y Turbany, 2007, pág. 130)

En este problema, se analiza una tabla 2x2, en la que existe una asociación entre las variables media-alta (coeficiente Phi de Pearson = -0,592), que es estadísticamente significativa ($p=0,000$). El alumno tiene teorías previas al respecto que apoyan el tipo de asociación y la relación sería de tipo inversa pues cuanto mayor es el tiempo de residencia, disminuye el grado de ansiedad. La relación sería de tipo causal, ya que sería el tiempo de residencia el que ocasiona un bajo ó alto grado de ansiedad.

Se espera que los estudiantes calculen las cuatro medidas de asociación que conocen para las tablas rxc y las interpreten. Analizamos a continuación la respuesta

correcta esperada en cada una de estas medidas.

Solución prevista para el coeficiente Phi de Pearson.

Para calcular el valor Phi de Pearson (Ato y López, 1996), los estudiantes disponen de un programa, que proporciona el valor 0,592. El programa no distingue entre signos, así que el estudiante tiene que tomar la decisión del signo mediante la observación de la tabla, donde las celdas *b* y *c* tienen mayor frecuencia que las *a* y *d*. Por tanto la asociación es negativa, así como el signo del coeficiente. Interpretando el resultado, el estudiante tiene que concluir que, como el valor del coeficiente está cerca de -0,6, la asociación es bastante alta (moderada-alta), y la asociación es negativa (estudiando las frecuencias de las celdas).

Solución prevista para el Riesgo relativo.

También el valor del riesgo relativo (Ato y López, 1996) se calcula con el programa, que proporciona el valor 0,3418 para el riesgo relativo por filas y por columnas. Los dos valores del riesgo relativo coinciden, lo que implica que es igual de fácil predecir el tiempo de residencia si se sabe el grado de ansiedad que al contrario, lo que también se observa las celdas de la tabla, ya que son iguales las frecuencias de las celdas *b* y *c*. El estudiante tiene que interpretar el resultado y concluir que:

- Para el riesgo relativo por columnas: Al predecir el grado de ansiedad, conociendo el tiempo de residencia se reduce el error casi en un 40% (respecto a si no sabemos nada). Diríamos que es más probable un grado de ansiedad “alto” si el tiempo de residencia es menor.
- Para el riesgo relativo por filas: Al predecir el tiempo de residencia, conociendo el grado de ansiedad se reduce el error casi en un 40%, siendo más probable un tiempo de residencia “alto” si el grado de ansiedad es bajo.

Solución prevista para la razón de productos cruzados.

El programa proporciona el valor 0,0493 para la razón de productos cruzados (Ato y López, 1996). El estudiante tiene que llegar a la conclusión de que, como el valor es inferior al valor 1, hay una asociación inversa. El coeficiente sugiere que la razón entre el número de emigrantes con “menos tiempo” de residencia y “más tiempo” de residencia es menor cuando el grado de ansiedad es bajo, es decir: $36/81 < 81/9$ o lo que

es igual, entre los emigrantes recientes sólo 36 se sienten con baja ansiedad por cada 81 que tienen alta ansiedad. Y respecto a los emigrantes que ya llevan bastante tiempo, 81 se encuentran con baja ansiedad frente a 9 que la tienen alta.

6.3.3. ANÁLISIS DEL PROBLEMA C: MEDIDAS DE ASOCIACIÓN EN TABLAS RXC

Problema C. Se está realizando un estudio sobre la población de estudiantes que hicieron la prueba de selectividad en el curso 1978-1979. Dos de las variables registradas fueron “convocatoria” (junio-septiembre) y “tipo de estudios” por el que se inclinaba el estudiante (grado medio, facultades, escuelas técnicas). Los resultados encontrados en la Universidad Complutense fueron los expuestos en la tabla siguiente:

	Escuelas Técnicas	Facultades	Grado Medio
Septiembre	200	500	800
Junio	1500	1300	700

Calcule e interprete las medidas de asociación para tablas rxc
(Adaptado de San Martín y Pardo, 1989, pág. 449)

En este problema, se plantea el cálculo de las medidas de asociación en una tabla 2x3, donde existe una asociación entre las variables media-baja (coeficiente V de Cramer= 0,364), por lo que es estadísticamente significativa ($p=0,000$). El alumno no tiene teorías previas al respecto y la relación sería de tipo tercera variable pues es posible que se produzca la relación por alguna variable desconocida.

Solución prevista para el cálculo del estadístico Chi-cuadrado.

Para resolver el problema, el estudiante ha de calcular el estadístico Chi-cuadrado, especificando sus grados de libertad que son: $(2-1) \times (3-1)=2$. Dispone de un programa para el cálculo de los componentes de la Chi-cuadrado y su valor total, con ayuda del cual calcula el valor Chi cuadrado, que toma el valor 662,31.

Solución prevista para el coeficiente de contingencia de Pearson.

Para calcular el coeficiente de contingencia de Pearson (Ruiz-Maya, Martín, Montero y Uriz, 1995), los estudiantes pueden utilizar el programa, con el cual obtienen el valor 0,342, así como el valor máximo, que es 0,7071. El estudiante tiene que interpretar este resultado y concluir que, dado que el valor del coeficiente está un poco por debajo de la mitad de su valor máximo, se puede considerar que hay una dependencia moderada.

Solución prevista para la Lambda de Goodman y Kruskal.

Para calcular el valor de Lambda de Goodman y Kruskal (Ruiz-Maya, Martín, Montero y Uriz, 1995; Batanero y Díaz, 2008), los estudiantes tienen que introducir en el programa el número de columnas y número de filas, y obtendrán el valor 0,067, es decir, la reducción del error de predicción es sólo del 6,7%.

El estudiante tiene que llegar a la conclusión de que el error apenas se ve reducido al predecir el valor de la variable dependiente *Y*, conocido el valor de la variable independiente *X*. En este caso significa que es difícil predecir de qué facultad es el estudiante si se sabe en que convocatoria aprobó. Además, si cambiamos filas y columnas, se obtiene un valor Lambda de Goodman y Kruskal = 0,167. Es algo más fácil (pero no demasiado, pues solo se reduce el error el 16,7%) deducir en qué convocatoria aprueba el estudiante si se sabe qué carrera estudia.

Solución prevista para la V de Cramer.

Para calcular el valor *V* de Cramer (Ruiz-Maya, Martín, Montero y Uriz, 1995), los estudiantes obtienen el valor 0,364 usando el programa Excel. Al interpretar el resultado han de concluir que está un poco por debajo de la mitad de su valor posible; por lo que también este coeficiente indica que hay una dependencia moderada.

6.3.4. ANÁLISIS DEL PROBLEMA D: CÁLCULO DE PROBABILIDADES Y JUICIOS DE ASOCIACIÓN

Problema D. Un grupo de 200 personas aquejadas de ansiedad fue dividido aleatoriamente en dos subgrupos. Al primer grupo se ofreció unas píldoras realmente efectivas para que tomaran tres al día, y al otro se ofreció un placebo (medicamento sin efecto). Al cabo de un mes fueron interrogados sobre la eficacia de las pastillas tomadas, con el siguiente resultado:

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

- a. Preguntado un paciente, dice que su ansiedad casi ha desaparecido; está encantado con el tratamiento. ¿Cuál es la probabilidad de que este paciente haya tomado el placebo?
- b. ¿Hay asociación entre el tipo de tratamiento (medicamento o placebo) y el efecto producido (la ansiedad disminuye o no)? ¿O son las variables independientes?
- c. Indica cómo has llegado a esta conclusión (puedes usar el método que prefieras)

En este problema, se plantean varias 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 mejore la ansiedad y la relación sería de tipo interdependencia. Este

problema fue resuelto sin ayuda del ordenador.

Solución prevista para el apartado (a). Probabilidad condicional.

Se plantea al estudiante el cálculo de una probabilidad condicional a partir de una tabla de contingencia, una tarea que ha sido propuesta en otras investigaciones (como las de Díaz, 2007 o Contreras, 2011), en las que se describen errores frecuentes, como la confusión entre probabilidad condicional y conjunta o el intercambio de los términos en la probabilidad condicional. Para resolver este apartado el estudiante ha de leer e interpretar los valores de la tabla, además de saber identificar el valor del numerador y del denominador de dicha probabilidad. La solución correcta en este apartado sería: $P(\text{Placebo}|\text{ansiedad ha disminuido}) = \frac{96}{(50 + 96)} = \frac{96}{146} = 0,657$.

Solución prevista para los apartados (b) y (c). Dependencia ó independencia.

El estudiante debe hacer un juicio de asociación, a partir de los datos dados en una tabla de contingencia 2x2 (Batanero y Díaz, 2008), es decir, es una tarea similar a las propuestas en el cuestionario inicial descrito en el Capítulo 3. 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.

6.3.5. ANÁLISIS DEL PROBLEMA E: JUICIO DE ASOCIACIÓN Y JUSTIFICACIÓN

Problema E. En un estudio se analiza la relación entre la dieta pobre en calcio y el desarrollo de osteoporosis. Se obtienen los siguientes resultados

Dieta	Desarrolla osteoporosis	
	Si	No
Pobre en calcio	58	62
Rica en calcio	22	258

- ¿Hay asociación entre el tipo de dieta (rica o pobre en calcio) y el desarrollo de osteoporosis?
- ¿Es la relación directa o inversa?
- ¿Es fuerte o moderada?
- Utiliza algún coeficiente de asociación que recuerdes para justificar tus respuestas
- Si no recuerdas los coeficientes, da la justificación usando cualquier propiedad que hayas aprendido en los temas de tablas de contingencia

En este problema, se analiza una tabla 2x2, donde existe una asociación media entre las variables (coeficiente Phi de Pearson=0,464) que es estadísticamente significativa ($p=0,000$). El alumno tiene teorías previas al respecto y la relación sería de tipo causal, pues sería el tipo de dieta el que influye en el desarrollo de la osteoporosis. Este problema se resolvió sin ayuda del ordenador.

Solución prevista para el apartado (a). Dependencia ó independencia.

De nuevo se pide en este apartado realizar un juicio de asociación (Batanero y Díaz, 2008), que, como en el problema anterior, puede llevarse a cabo con diversos procedimientos. El estudiante podría realizar una comparación entre las frecuencias esperadas en caso de independencia con las frecuencias observadas, concluyendo que son diferentes; esto indica que hay relación entre las variables, pues las diferencias son importantes.

El estudiante podría comparar las frecuencias relativas condicionales por filas o por columnas o las frecuencias relativas condicionales con las marginales, llegando a la misma conclusión. Otra posible solución parcialmente correcta sería que el estudiante justificase el tipo de dependencia según el tipo de celdas en donde se concentra la mayor frecuencia de sujetos del problema según Inhelder y Piaget (1955), es decir, que comparase los valores de las celdas que informan de una dependencia directa (58 y 258), con los de las celdas que corresponden a una dependencia inversa (62 y 22). Concluiría que hay asociación, ya que la mayor parte de los sujetos del experimento están en las celdas de dependencia directa, pues tenemos 316 casos de 400 (las tres

cuartas partes). No sería totalmente correcta pues esta estrategia es válida únicamente cuando las frecuencias marginales son iguales.

Solución prevista para el apartado (b). Dependencia directa ó inversa.

Una vez identificada la dependencia, el estudiante podría comparar los valores de las celdas que informan de una dependencia directa (58 y 258), con los de las celdas que informan de una dependencia inversa (62 y 22) según Inhelder y Piaget (1955). Con esta estrategia concluiría que la relación es directa, ya que las celdas de la diagonal donde están los sujetos, son las celdas que informan de una dependencia directa. Por tanto, si la dieta es pobre en calcio hay mayor osteoporosis. Podría también calcular alguno de los coeficientes de asociación para tablas 2x2 e interpretar su valor para deducir el signo de la asociación.

Solución prevista para los apartados (c) y (d). Intensidad

El estudiante debe calcular un coeficiente de asociación (Ruiz-Maya, Martín, Montero y Uriz, 1995; Ato y López, 1996; Pardo y San Martín, 1998), y basarse en su valor para concluir si la dependencia es fuerte o moderada. El estudiante puede calcular cualquiera de los coeficientes estudiados, que son:

- Phi de Pearson = 0,464. Ya que el valor Phi de Pearson está por debajo del 0,5 se considera que la dependencia es moderada, pues este coeficiente puede tomar un valor hasta 1. Quiere decir que, además de la dieta habrá otros factores que influyan en la osteoporosis.
- El valor del riesgo relativo por columnas igual a 3,741 indica que es 3,7 veces más probable tener dieta pobre en calcio si se tiene osteoporosis que si no se tiene. El riesgo relativo por filas = 6,1515; indica que es 6,15 veces más probable tener osteoporosis con dieta pobre en calcio que con dieta normal.
- Razón de productos cruzados = 10,97. Al ser mayor que 1, la dependencia es directa. El cociente entre casos con y sin osteoporosis es 10 veces mayor si la dieta es pobre en calcio.

Solución prevista para el apartado (e). Justificación sin usar un coeficiente.

Una posible solución, es que el estudiante compare las frecuencias observadas con las esperadas, concluyendo que son muy diferentes; esto indica que la relación entre

las variables es moderada, pues las diferencias existentes no son exageradamente altas (Tabla 6.3.5.1).

Tabla 6.3.5.1. Tabla de frecuencias esperadas

Dieta	Desarrolla osteoporosis	
	Si	No
Rica en calcio	24	96
Pobre en calcio	56	224

6.4. SINTESIS DE LAS EVALUACIONES MEDIANTE PROBLEMAS ABIERTOS

Como síntesis del análisis realizado, mostramos en la Tabla 6.4.1 los objetos matemáticos incluidos en los problemas abiertos usados en la evaluación. Observamos que se han tenido en cuenta muchos de los objetos incluidos en la enseñanza, lo que dotará de validez de contenido al proceso de evaluación

Tabla 6.4.1. Objetos considerados en los problemas abiertos

Tipo	Objetos matemáticos	Problema
Sit-pro.	P1. Resumir la información obtenida de un conjunto de observaciones bivariantes de modo que se visualicen posibles relaciones entre las variables.	Todos
	P2. Cálculo de probabilidades asociadas a valores específicos de una o las dos variables.	D
	P3. Contraste de homogeneidad.	A
	P4. Contraste de independencia.	D, E
	P5. Definir una medida de la intensidad de asociación entre las variables.	B, C
Leng.	R1. Lenguaje verbal.	Todos
	R2. Lenguaje simbólico.	Todos
	R3. Tablas (de frecuencia y de contingencia).	Todos
Procedimientos	PR1. Lectura e interpretación de la tabla	Todos
	PR2. Cálculo de probabilidades simples, compuestas y condicionales	D, E
	PR3. Cálculo de frecuencias condicionales y marginales	A, B, C, D
	PR4. Manejo del software.	A, B, C
	PR5. Fórmula de cálculo de las frecuencias esperadas	Todos
	PR6. Interpretación de medidas en tablas 2x2	B, E
	PR7. Interpretación de medidas en tablas rxc	C
	PR8. Interpretación de medidas basadas en la reducción del error	C
	PR9. Formulación de hipótesis (contraste de independencia)	D, E
	PR10. Formulación de hipótesis (contraste de homogeneidad)	A
	PR11. Interpretación de resultados del test χ^2 cuadrado obtenido con el ordenador	B, C
	PR12. Completar un contraste de hipótesis tomando una decisión	A
	PR13. Interpretar el resultado de un contraste en el contexto de los problemas	A, B, C
	PR14. Calcular frecuencias relativas marginales a partir de las absolutas marginales	D, E
	PR15. Independencia si coinciden frecuencias condicionales por columnas ó filas	Todos
	PR16. Cálculo de los grados de libertad para realizar un test χ^2 cuadrado	A, B, C
	PR17. Si el valor Chi cuadrado es estadísticamente significativo	A
	PR18. Manejo de tablas de distribuciones de probabilidad	A
	PR19. Calcular el parámetro ó grados de libertad	A, B, C
	PR20. Que valor p indica más asociación	A
	PR21. Aceptar/rechazar la hipótesis nula basándonos en el valor p	A, D, E

	PR22. Aceptar/rechazar la hipótesis nula basándonos en el estadístico	A
	PR23. Selección de un medidor de la asociación	B, C, E
	PR24. Comprobar las propiedades de independencia	A, D, E
	PR25. Cálculo del estadístico de contraste	A
Conceptos	C1. Variable simple / doble valor.	Todos
	C2. Tablas rxc	A, C
	C3. Frecuencia absoluta doble y marginal	Todos
	C4. Frecuencia relativa marginal.	Todos
	C5. Frecuencia relativa condicional.	Todos
	C6. Frecuencia relativa doble	Todos
	C7. Distribución marginal.	Todos
	C8. Distribución condicional.	Todos
	C9. Probabilidad conjunta y condicional	D
	C10. Independencia /Asociación	D, E
	C11. Dependencia aleatoria y dependencia funcional	Todos
	C12. Asociación inversa, directa.	B, E
	C13. Intensidad de la asociación.	B, C, E
	C14. Frecuencias esperadas	Todos
	C15. Estadístico Chi cuadrado, parámetro, distribución	A, B, C
	C16. Grados de libertad	A, B, C
	C17. Coeficiente Phi, riesgo relativo, razón de productos cruzados	B, E
	C18. Coeficiente C y V (medidas simétricas, tablas rxc)	C
	C19. Coeficiente Lambda (medidas asimétricas)	C
	C20. Significación	A
	C21. p valor	A
	C22. Hipótesis alternativa y nula	A
Propiedades	PP1. Igualdad frecuencias relativas condicionales en caso de independencia	D, E
	PP2. Igualdad frecuencias relativas condicionales y marginales en independencia	D, E
	PP3. Celdas que indican signo de asociación en la tabla 2x2	C, D, E
	PP4. Frecuencias esperadas en caso de independencia	D, E
	PP5. Relación entre asociación y causalidad	Todos
	PP7. Relación entre valor p en el contraste Chi cuadrado y asociación	A
	PP8. Diferencias entre el significado y aplicación de las medidas de asociación	A
	PP9. Propiedades de independencia	Todos
	PP10. Cuando un valor chi cuadrado es estadísticamente significativo	D, E
	PP11. Un valor $p < 0,05$ es estadísticamente significativo si no se dice otro valor	A
	PP12. Condiciones de aplicación del estadístico Chi cuadrado	A
	PP13. Relación entre el riesgo relativo por filas y columnas	A, B, C
	PP14. Relación entre probabilidad y tabla de contingencia	B
	Arg.	A1. Razonamientos informales.
A2. Comprobación de ejemplos y contraejemplos, sin pretensión de generalizar.		D, E
A3. Deductivos		Todos

6.5. METODOLOGÍA DE ANÁLISIS DE LOS ÍTEMS ABIERTOS

Como se ha señalado, en la enseñanza participaron 104 estudiantes de Psicología de la Universidad de Granada, divididos en dos grupos, que constaban de 51 estudiantes el primero (8 hombres y 43 mujeres) y 53 estudiantes el segundo (16 hombres y 37 mujeres). Todos ellos eran estudiantes de segundo año y cursaban la asignatura de “*Técnicas de análisis en la investigación psicológica*”. La mayoría tenía una edad de entre 19 y 20 años.

Las soluciones a los problemas abiertos se recogieron en las sesiones prácticas y de evaluación de la citada asignatura. El número de participantes en las tablas

presentadas en este capítulo puede variar de unos problemas a otros, ya que cada uno se realizó en momentos diferentes, y es posible que un estudiante asistiera un día y otro no, pues la asistencia no es obligatoria.

Aún así, como veremos hay una fuerte participación, lo que indica que se tuvo éxito al motivar a los estudiantes por este tema. El estudiante trabajó individualmente, sin poder consultar a sus compañeros o al profesor; debía ser capaz de interpretar cada enunciado y aplicar lo aprendido en la resolución del problema. La parte más compleja es la interpretación de los resultados del trabajo matemático en el contexto; es precisamente esta última parte la que mostrará un aprendizaje más significativo del estudiante. En los tres primeros problemas (A, B, C) se les dejó utilizar el ordenador, por lo cuál, aunque no necesitaban memorizar los procedimientos, aunque deberían recordar el conocimiento a aplicar y realizar en forma correcta esta aplicación.

Recogidos los datos se procedió a realizar un análisis de contenido, clasificando las diferentes respuestas en cada apartado, mediante un proceso inductivo y cíclico, hasta llegar a un número suficiente de categorías para mostrar la diversidad de soluciones. Al igual que en el Capítulo 3, se aplicó un análisis de contenido, que nos permite pasar del texto a una serie de variables estadísticas de las cuáles posteriormente se realizan inferencias (Ghiglione y Matalón, 1991). Siguiendo a estos autores hemos considerado el *análisis de contenido temático*, utilizando nuestro conocimiento previo sobre el tema, para resumir el contenido del texto y definir las categorías de análisis y también nos hemos apoyado en la revisión bibliográfica presentada en el Capítulo 2.

La primera operación fue la lectura atenta de las producciones de los estudiantes, buscando las componentes relacionadas con cada uno de los apartados de cada problema. Las respuestas de todos los estudiantes fueron transcritas en un fichero de texto, para poderlas comparar, asignando al estudiante su número y a cada respuesta un código correspondiente al problema y apartado. Las respuestas de todos los estudiantes al mismo problema y apartado fueron clasificadas y comparadas entre sí, así como con los antecedentes. Tanto los directores, como el autor de la memoria revisaron varias veces la codificación, hasta llegar a un acuerdo sobre las categorías que se utilizarán en el estudio. En lo que sigue se exponen los resultados de este proceso.

6.6. RESULTADOS EN EL PROBLEMA A: CONTRASTE DE HOMOGENEIDAD

Problema A. Para un experimento sobre memoria a largo plazo un psicólogo selecciona 3 palabras y decide evaluar el componente emocional de las mismas con una muestra aleatoria. Presenta, independientemente, cada palabra a 100 sujetos y registra si la palabra es percibida como emocionalmente positiva (agradable, placentera, etc.), negativa (desagradable, displacentera, etc.) o neutra. A la vista de los resultados obtenidos, ¿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.

(Adaptado de Pardo y San Martín, 1998, pág. 554)

En este problema se pide a los estudiantes seguir los pasos pertinentes en un ejercicio típico de un contraste de homogeneidad. Se le proporciona una tabla 3x3, y comienza con el planteamiento de las hipótesis nula y alternativa adecuadas en un contraste de este tipo. Posteriormente, el estudiante tiene que calcular el estadístico de contraste, y la probabilidad asociada al valor obtenido. Finalmente el último paso es la interpretación de los resultados en el contexto del problema. Los alumnos podían consultar sus apuntes. Encontramos cuatro apartados, que estudiaremos de forma independiente en lo que sigue. Los resultados del análisis se han publicado en Cañadas, Batanero, Díaz y Roa (2012).

6.6.1. FORMULACIÓN DE HIPÓTESIS

El primer paso para resolver un problema de contraste de hipótesis es establecer dichas hipótesis, donde los estudiantes con frecuencia tienen algunas dificultades, pues algunos confunden las hipótesis nula y alternativa o plantean las hipótesis de forma que no cubren el espacio muestral o utilizando el estadístico y no el parámetro (Vallecillos, 1994). Hemos encontrado dos tipos diferenciados de respuestas correctas, cinco de respuestas parcialmente correctas y ocho de respuestas incorrectas.

Respuestas correctas:

Consideramos correctas las respuestas de los estudiantes que han interpretado correctamente el enunciado y establecen de modo adecuado las hipótesis nulas y alternativa de un contraste de homogeneidad. Hemos encontrado dos variantes:

RC.1. El estudiante usa una notación adecuada tanto para la hipótesis nula como alternativa, y hace referencia explícita al parámetro “proporción poblacional” en forma simbólica al establecer las hipótesis, que son elegidas en forma correcta (Lipschutz y Schiller, 1999; Ruiz-Maya y Martín, 2002). A continuación mostramos un caso en que el estudiante ha podido interpretar el enunciado y comprender el tipo de hipótesis que se plantearían en un contraste de homogeneidad, pues no se limita a copiar la definición presentada en los apuntes, sino que es capaz de aplicarla al caso concreto presentado. El estudiante usa los conceptos de población y subpoblación, así como el de proporción (parámetro poblacional). La notación empleada indica que diferencia tanto la población como el elemento de la misma:

$H_0: p_{1j} = p_{2j} = p_{3j}$
 $H_1: \text{Alguna subpoblación es diferente}$ (Estudiante 51).

RC.2. En esta segunda formulación se hace referencia al concepto de variable y distribución de la variable (en vez de hacer referencia al parámetro poblacional) y no se usa la expresión simbólica de dicho parámetro. Las hipótesis nula y alternativa están correctamente formuladas, por lo que el estudiante diferencia entre estas dos hipótesis, a pesar de ser un punto difícil, de acuerdo a Vallecillos (1994) y Chow (1996). Por ejemplo, hemos encontrado el siguiente caso, donde también se usan los conceptos de población y subpoblación, distribución y variable:

$H_0: \text{Todas las subpoblaciones tienen idéntica distribución para la variable } Y.$
 $H_1: \text{La distribución de la variable } Y \text{ en alguna de estas subpoblaciones es diferente}$ (Estudiante 53).

Respuestas parcialmente correctas:

Hemos considerado parcialmente correctas aquellas respuestas en que los estudiantes eligen una hipótesis nula relacionada con el concepto de homogeneidad de poblaciones, y como alternativa la complementaria, siempre haciendo referencia a las tres poblaciones. No son completamente correctas por haber un fallo en la elección del parámetro o su simbolización o en la verbalización de las hipótesis. Hemos encontrado las siguientes variantes:

RP.1. Aunque el estudiante elige correctamente la hipótesis nula (como de igualdad de poblaciones) y la alternativa (como de diferencia entre alguna de ellas); hace referencia (al menos en forma simbólica) a la media poblacional, en vez de usar la proporción poblacional. Por tanto presenta una confusión sobre el tipo de variable en estudio (cualitativa) a la que no puede aplicarse el concepto de media. No hemos encontrado este error descrito en las investigaciones previas. Un ejemplo de esta categoría es el siguiente, en que el estudiante usa una notación correcta y los conceptos de media y población:

$H_0: \mu_1 = \mu_2 = \mu_3$
 $H_1: \mu_n \neq \mu_m$, para algún n, m perteneciente $\{1, 2, 3\}$ ” (Estudiante 73).

RP.2. El estudiante establece como hipótesis nula con un lenguaje poco preciso; generalmente no utiliza notación simbólica, por lo que es difícil deducir si está o no estableciendo las hipótesis en función del parámetro o del estadístico. Aparentemente comprende la hipótesis, pero no la expresa totalmente de forma correcta. Harradine, Batanero y Rossman (2011), indican que el concepto de distribución muestral es mucho más abstracto que los de distribución de una población o de una muestra, lo que puede causar dificultad al expresar las hipótesis. Un ejemplo se reproduce a continuación:

H_0 : Las tres palabras tienen el mismo componente emocional.
 H_1 : Las palabras tienen distintos componentes emocionales” (Estudiante 71).

RP.3. Se intercambian la hipótesis nula y alternativa y además el lenguaje es poco preciso. Tanto Vallecillos (1994) como Chow (1996) resaltan este error habitual entre estudiantes e investigadores, quienes además de confundir entre sí estas hipótesis, también las intercambian las hipótesis de investigación y las hipótesis estadísticas. Un caso es el siguiente, donde el estudiante, tendría que haber intercambiado las hipótesis, ya que decir que son homogéneas es lo mismo que decir que las palabras no influyen en las emociones. La falta de precisión no deja entrever qué conceptos está utilizando el estudiante:

H_0 : Las palabras presentadas influyen en las emociones.
 H_1 : Las palabras presentadas no influyen en las emociones” (Estudiante 19).

RP.4. Aunque el estudiante elige correctamente la hipótesis nula (todas las subpoblaciones tienen idéntica distribución para la variable Y), la alternativa no es correcta, ya que se exige diferencia entre alguna de ellas, pero no entre todas. Además la

forma simbólica no es la más adecuada, ya que no es usual. En la investigación de Vera, Díaz y Batanero (2011), se menciona que uno de los errores encontrados es que las dos hipótesis no cubren el espacio paramétrico; esto nos ocurre en este planteamiento. A continuación se muestra un caso:

*"H₀: Componente emocional₁ = Componente emocional₂ = Componente emocional₃
H₁: CE₁ ≠ CE₂ ≠ CE₃"* (Estudiante 24).

RP.5. El estudiante plantea correctamente la hipótesis nula, haciendo referencia a la distribución, en vez de a la variable; pero la hipótesis alternativa no sería correcta, ya que tendría que exigirse tan sólo la diferencia entre alguna de las tres proporciones y no entre todas ellas. Es decir, se repite el error de la respuesta RP4, por lo que las dos hipótesis no cubren el espacio paramétrico (error encontrado por Vera, Díaz y Batanero, 2011). Un ejemplo lo vemos a continuación, en el que el estudiante utiliza los conceptos de distribución y proporción (implícitamente pensamos que se refiere al parámetro, esto es, la proporción poblacional):

*"H₀: Las palabras 1,2 y 3 tienen la misma distribución.
H₁: Ninguna proporción es igual"* (Estudiante 41).

Respuestas incorrectas:

RI.1. El estudiante propone el valor del estadístico Chi-cuadrado en la hipótesis nula, y además no da ninguna hipótesis alternativa. En consecuencia, este estudiante confunde parámetro y estadístico, error encontrado y descrito en otras investigaciones (Vallecillos y Batanero, 1997; Castro Sotos et al., 2007). Schuyten (1991) también indica que no es sencilla para los estudiantes la diferencia entre el parámetro y el estadístico, debido a que el mismo concepto (por ejemplo media) se está usando a dos niveles diferentes: como estadístico (en la muestra dada) y como parámetro (en la población) y los estudiantes no llegan a diferenciar ambos planos. En esta respuesta se observa que el estudiante no entiende de forma correcta conceptos clave de un contraste de hipótesis, este tema aparece reflejado en otras investigaciones (Falk, 1986; Vallecillos y Batanero, 1997). Un ejemplo se incluye a continuación:

"H₀: 10,81" (Estudiante 5).

RI.2. El estudiante plantea las hipótesis correctas de un contraste de independencia, en lugar de las hipótesis de un contraste de homogeneidad. Aunque este error es razonable, puesto que los dos contrastes se realizan a partir de las tablas de contingencia y mediante el estadístico Chi-cuadrado, está confundiendo estos dos tipos

de contrastes y las situaciones en que uno y otro son aplicables. Además no discrimina la hipótesis nula de la alternativa, error que, como hemos indicado, ha sido descrito por otros investigadores (Vallecillos, 1994; Chow, 1996). Un ejemplo sería el siguiente, en que el estudiante usa los conceptos de variable, asociación e independencia:

*"H₀: Las variables están asociadas.
H₁: Las variables son independientes"* (Estudiante 62).

RI.3. El estudiante intercambia las hipótesis alternativa y nula, error ya comentado en otras respuestas. Por otro lado, en la hipótesis nula tendría que suponer únicamente que alguna variable es diferente, en lugar de exigir que todas lo sean. En esta estrategia las dos hipótesis no cubren el espacio paramétrico (Vera, Díaz y Batanero, 2011). A continuación mostramos un resultado de este tipo, donde se hace referencia al término variable; puesto que la diferencia de variables aleatorias implica la diferencia de su distribución aceptamos este uso:

*"H₀: X₁ ≠ X₂ ≠ X₃
H₁: X₁ = X₂ = X₃"* (Estudiante 64).

RI.4. El estudiante plantea un contraste de asociación (y no de homogeneidad), por lo que confunde estos dos tipos de contraste; además indica que la asociación se da entre filas o entre columnas de la tabla, en lugar de entre las filas y columnas simultáneamente. Por lo tanto, está confundiendo los valores que puede tomar una variable, con las distintas variables, al indicar que cada valor de una variable, es una variable diferente. White (1980) describe este error y supone que se debe a la dificultad que tienen algunos investigadores para identificar la población bajo estudio al aplicar la inferencia, coincidiendo también con Hagod (1970). Un ejemplo sería el que aparece seguidamente y donde el estudiante usa los conceptos variable y asociación:

*"H₀: Las variables entre columnas no están asociadas.
H₁: Las variables entre filas si están asociadas"* (Estudiante 65).

RI.5. El estudiante está confundiendo la variable bajo estudio (componente emocional) con las poblaciones de sujetos donde se estudia esta variable (datos tomados sobre las palabras). Además la hipótesis alternativa exige de nuevo la diferencia de todas las subpoblaciones, exigencia que, como se ha discutido, es excesiva. En consecuencia, como en los casos anteriores, en esta estrategia las dos hipótesis no cubren el espacio paramétrico (Vera, Díaz y Batanero, 2011). Una respuesta de este tipo sería:

“ $H_0: W_1 = W_2 = W_3$ ($W=$ palabra).
 $H_1: W_1 \neq W_2 \neq W_3$ ” (Estudiante 83).

RI.6. Usa dos poblaciones, en lugar de tres; posiblemente el estudiante está confundiendo el número de variables con el número de subpoblaciones que se comparan. Se observa por consiguiente una confusión entre variable y población, ya descrita y una dificultad para identificar la población bajo estudio al aplicar la inferencia (Hagod, 1970; White, 1980). Un ejemplo sería el que presentamos a continuación:

“ $H_0: p_{1j} = p_{2j}$
 $H_1: \text{son distintos}$ ” (Estudiante 92).

RI.7. El estudiante plantea bien la hipótesis nula, pero la hipótesis alternativa sólo considera que sean distintas las dos primeras distribuciones. Puesto que la hipótesis alternativa ha de exigir que alguna (sin especificar cuál) población al menos sea diferente, como en otros casos esta hipótesis está mal formulada. En esta estrategia las dos hipótesis no cubren el espacio paramétrico (Vera, Díaz y Batanero, 2011). A continuación mostramos una respuesta en esta categoría:

“ $H_0: p_{1j} = p_{2j} = p_{3j}$
 $H_1: p_{1j} \neq p_{2j} = p_{3j}$ ” (Estudiante 95).

En la formulación de las hipótesis nulas de un contraste de homogeneidad, podemos observar en la Tabla 6.6.1.1 como 43 estudiantes (46,6%) plantean correctamente las hipótesis, 29 estudiantes (31,6%) plantean parcialmente correcta, y 13 estudiantes (14,2%) las plantean de forma incorrecta. Los resultados son mejores que los de Vallecillos (1994) y Vera, Díaz y Batanero, donde 26% y 33% de los participantes respectivamente plantean correctamente las hipótesis. En consecuencia, estos estudiantes tienen unos resultados razonables en el planteamiento de las hipótesis. Hacemos notar que en los apuntes entregados a los estudiantes muchas de las reglas que hemos comentado (como que las hipótesis han de cubrir el espacio paramétrico) sólo se mencionan en forma implícita.

Las respuestas más frecuentes son las RC.1 y RC.2, en la que plantean las hipótesis de un contraste de homogeneidad de forma correcta una de forma simbólica y otra verbalmente. Las dos siguientes estrategias que destacan son parcialmente correctas: (RP.1), donde hace referencia al parámetro media muestral, en vez de usar la media poblacional, es decir, el único error de frecuencia a considerar es la confusión entre estadístico y parámetro, señalada por Schuyten (1991) y Vallecillos (1994).

Tabla 6.6.1.1. Frecuencias y porcentajes de respuestas en el problema A, apartado a.

Respuesta		Frecuencia	Porcentaje	% H ₀ correcta	% H ₁ correcta
Correcta	RC.1	29	31,4	31,4	31,4
	RC.2	14	15,2	15,2	15,2
Parcialmente correcta	RP.1	10	10,9		
	RP.2	7	7,6		
	RP.3	1	1,1		
	RP.4	2	2,2	2,2	
	RP.5	10	10,9	10,9	
Incorrecta	RI.1	1	1,1		
	RI.2	1	1,1		
	RI.3	4	4,3		
	RI.4	2	2,2		
	RI.5	1	1,1		
	RI.6	2	2,2		
	RI.7	1	1,1		
No responde		7	7,6		
Total		92	100	59,7	46,6

Es también relativamente frecuente la respuesta RP.5 que en la hipótesis alternativa se exige que todas las poblaciones sean diferentes. Si se suman a esta categoría las RI.3, RI.5, RI.6, casi un 20% de estudiantes plantean hipótesis que no cubren el espacio paramétrico, error también encontrado en el trabajo de Vera, Díaz y Batanero (2011).

Por último, según nuestros datos, la formulación de la hipótesis nula ha resultado más fácil para nuestros estudiantes que la formulación de la hipótesis alternativa (60% frente al 46,6%). Posiblemente la dificultad estriba en el hecho de que las hipótesis se refieren a tres subpoblaciones, mientras que en los contraste estudiados anteriormente (por ejemplo, el contraste t) se comparan solamente una o dos poblaciones; ello puede haber influido en no saber formular correctamente la hipótesis alternativa.

Finalmente resaltamos el hecho de que aproximadamente el 47% de estudiantes plantearon correctamente las dos hipótesis, resultados bastante buenos, comparados con los obtenidos en las investigaciones previas sobre otros tipos de contrastes, pero aún insuficientes, ya que el planteamiento correcto de las hipótesis no llega a la mitad de los estudiantes de la muestra.

6.6.2. ESTADÍSTICO DE CONTRASTE Y PROBABILIDAD ASOCIADA

Una vez planteadas las hipótesis, el estudiante debe usar el programa para obtener el valor Chi-cuadrado. A continuación ha de dar los grados de libertad y el programa, entonces, le proporciona el valor p . Las respuestas encontradas son las

siguientes:

Respuestas correctas:

RC.1. El estudiante usa una notación adecuada tanto para el estadístico de contraste, como para el valor p , y obtiene los resultados correctos para ambos casos. Para la obtención correcta del estadístico, tiene que introducir las frecuencias absolutas de la tabla de contingencia en el programa, construyendo de esta forma la tabla. Una vez que el estudiante determina correctamente los grados de libertad y los introduce en el programa, el cual proporciona el valor p correcto (Batanero y Díaz, 2008). El estudiante ha podido reconocer de forma correcta el estadístico de contraste como el valor Chi-cuadrado, y la probabilidad asociada al valor obtenido como el valor p . Un resultado de este tipo es:

“Chi-cuadrado experimental = 10,81; p-valor = 0,029” (Estudiante 71).

Respuestas parcialmente correctas:

RP.1. En esta solución el estudiante calcula bien el valor del estadístico Chi-cuadrado, lo que quiere decir que construye de forma correcta la tabla de contingencia en el programa, pero el valor de p es incorrecto. El número de grados de libertad también es correcto aparentemente, pero la transcripción del programa al papel del valor p se realizó de forma incorrecta, dando 0,021, en lugar de 0,029; en esta solución sólo varía una cifra decimal y únicamente un estudiante dio esta respuesta. Este estudiante aunque proporciona una solución parcialmente correcta, identificó los valores que se le pedían y los calculó de forma correcta. A continuación mostramos el único resultado de este tipo:

“Chi-cuadrado experimental = 10,81; p-valor = 0,021” (Estudiante 7).

Respuestas incorrectas:

RI.1. El estudiante calcula correctamente el valor Chi-cuadrado, lo que implica una correcta construcción de la tabla de contingencia en el programa, pero no recuerda los grados de libertad ó los calcula mal, ya que considera otros para el cálculo del valor p (en el ejemplo se considera 1 g. l., pero se han encontrado respuestas con 2 g. l.), en lugar de cuatro. Por tanto obtiene un valor p incorrecto al no calcular bien los grados de libertad (Batanero y Díaz, 2008).

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 . Ya Walker (1940) reflexionó sobre la dificultad del concepto de grados de libertad, que tiene una importancia central en la moderna teoría de la estadística, pero al que se presta poca atención. Dicho concepto se usa en varias distribuciones muestrales, cuya probabilidad cambia con los grados de libertad, que aparecen como parámetro en la distribución. Como consecuencia, un error al determinar los grados de libertad llevará a encontrar una probabilidad errónea, como ha resultado en esta respuesta. Un ejemplo se incluye a continuación

“Chi-cuadrado experimental = 10,8; $p = 0,001$ ” (Estudiante 36).

RI.2. El estudiante calcula correctamente el valor Chi cuadrado, mediante la introducción de forma correcta de las frecuencias absolutas y los grados de libertad (4 g.l.). Pero, en lugar de obtener el valor p con el programa, utiliza la tabla de la distribución Chi-cuadrado, donde para $\alpha=0,01$ y 4 g.l. el valor crítico es 13,28. Toma este valor crítico como si fuese el p -valor, manifestando un conflicto, ya que confunde estos dos conceptos. Por otro lado, no observa que da un valor de una probabilidad (es decir de p) mayor que uno lo cual es otro conflicto, también encontrado por Contreras (2011). En consecuencia, el estudiante confunde conceptos necesarios en los contrastes de hipótesis (Falk, 1986; Vallecillos y Batanero, 1997). Una respuesta de este tipo, sería:

“Chi-cuadrado experimental = 10,81: $p = 13,28$ ” (Estudiante 61).

RI.3. El estudiante realiza una construcción eficiente de la tabla de contingencia en el programa, lo que le proporciona un valor correcto del estadístico Chi-cuadrado. Por otro parte, no sabe ó reconoce, el valor de la probabilidad asociada al valor obtenido, que se le pide, es decir de p ; el conflicto posiblemente sea que no recuerda la fórmula de los grados de libertad, error que se encontró en otros trabajos (Alvarado, 2007; Olivo, 2008). Un ejemplo es:

“Chi-cuadrado experimental = 10,81” (Estudiante 56).

RI.4. El estudiante obtiene el estadístico de contraste, pero da el valor de alfa en vez de calcular el valor p que se le pide, es decir, el conflicto aparece en que probablemente no recuerda los grados de libertad o que no diferencia entre nivel de significación y p -valor. Hay una gran variedad de investigaciones que su estudio se centra en la comprensión del concepto de nivel de significación e indican la dificultad

de este concepto (Birnbaum, 1982; Vallecillos, 1994; Krauss y Wassner, 2002). Bakan (1966) también indica que muchos usos incorrectos de los contrastes de hipótesis tienen su origen en la interpretación incorrecta del concepto de nivel de significación. A continuación vemos un caso:

“*Chi-cuadrado experimental = 10,81; p = 0,01*” (Estudiante 83).

En la Tabla 6.6.2.1 tenemos las frecuencias y porcentajes de cada respuesta, observándose como 58 estudiantes (63,1%) reconocen correctamente estos valores (el estadístico Chi-cuadrado y el *p*-valor), y el resto no es capaz de determinarlos, a pesar de que tenían el programa de Excel a su disposición como ayuda. Las respuestas correctas son las más frecuentes, a las que hay que añadir el caso del estudiante que da una respuesta parcialmente correcta.

Resaltan entre las respuestas incorrectas (RI.1), en la que no calculan o no recuerdan los grados de libertad, error descrito por Alvarado (2007) y Olivo (2008); y (RI.3), donde el estudiante no reconoce, el valor de la probabilidad asociada al valor obtenido, que se pide, es decir de *p*, lo cual también puede ser debido a los grados de libertad. Los conflictos principales son los relacionados con el valor *p*, en el cálculo de los grados de libertad y en relacionar este símbolo con el valor de la probabilidad asociada al valor obtenido.

Tabla 6.6.2.1. Frecuencias y porcentajes de respuestas en el problema A, apartado b

Respuesta		Frecuencia		Porcentaje		% Chi-cuadrado correcto	% g.l.	% p-valor correcto
Correcta	RC.1	58	63,1	63,1	63,1	63,1	63,1	63,1
Parcialmente correcta	RP.1	1	1,1	1,1	1,1	1,1	1,1	
Incorrecta	RI.1	14	15,2	15,2				
	RI.2	6	6,5	6,5		6,5		
	RI.3	12	13	13,0				
	RI.4	1	1,1	1,1				
Total		92	100	100	70,7	63,1		

Todos los estudiantes han calculado bien el valor de Chi-cuadrado, sin embargo solamente el 70,7% calculan correctamente los grados de libertad, en consonancia con investigaciones anteriores ya comentadas. Por último, el cálculo del *p*-valor ha sido el más difícil, pues solo lo calcula correctamente el 63.1%.

6.6.3. TOMA DE DECISIÓN

Una solución posible en este apartado (que no requiera el uso de tablas estadísticas) sería razonar que como el p -valor = 0,029 > α = 0,01, entonces no podemos rechazar la hipótesis nula H_0 . El estudiante debe recordar la lógica del contraste de hipótesis y la forma en que se construyen la región de aceptación y rechazo, bien a partir del valor crítico o del p -valor. A continuación se describen y analizan los tipos de respuestas obtenidos.

Respuestas correctas:

RC.1.No podemos rechazar la hipótesis nula H_0 . Sería el estudiante que llega a una conclusión correcta, interpretando adecuadamente los resultados del contraste, mediante uno de los dos métodos enseñados: comparar el estadístico de contraste con el valor crítico de la tabla de la distribución Chi-cuadrado, ó comparar el valor p con el valor de alfa. Hemos aceptado como equivalentes diferentes formulaciones, tales como la siguiente:

“No podemos aceptar que las muestras sean diferentes” (Estudiante 59).

Respuestas incorrectas:

RI.1.Rechazamos la hipótesis nula H_0 . Serían los estudiantes que deciden rechazar la hipótesis mediante uno de los métodos dicho anteriormente. Al realizar un juicio de asociación (aceptar ó rechazar una hipótesis nula), basándonos en los resultados del contraste, se consideran dos tipos de error (Nortes Checa, 1993; Peña y Romo, 1997): *error tipo I*, rechazar una hipótesis nula que de hecho sea verdadera; y *error tipo II*, aceptar la hipótesis nula que realmente es falsa. En nuestro caso, los estudiantes tienen un conflicto relacionado con el *error tipo I*. Estos estudiantes no han comprendido la lógica del contraste de hipótesis, un punto difícil señalado por Harradine, Batanero y Rossman (2011).

Tabla 6.6.3.1. Frecuencias y porcentajes de respuestas en el problema A, apartado c

	Frecuencia	Porcentaje
RC.1	54	58,7
RI.1	30	31,3
No responde	8	8,7
Total	92	100

En la toma de la decisión, de aceptar o no la hipótesis nula, podemos observar en la Tabla 6.6.3.1 como 54 estudiantes (58,7%) toman la decisión correcta de no rechazar

la hipótesis nula, es decir la mayoría, aunque este punto fue difícil en investigaciones previas como la de Vallecillos (1994). Observamos que también cerca de la mitad de los estudiantes no llegan a tomar la decisión adecuada, por fallo en la comprensión de la lógica del contraste, error ya señalado por diferentes autores (Birnbaum, 1982; Vallecillos, 1994; Lecoutre, Lecoutre y Poitevineau, 2001; Haller y Kraus, 2002; Lecoutre, 2006).

6.6.4. INTERPRETACIÓN DE RESULTADOS

El último paso es la interpretación de los resultados en el contexto del problema, es decir, traducir los resultados estadísticos a términos de lo que busca el investigador. Como no podemos rechazar la hipótesis nula con un nivel de significación del 0,01, esperamos que los estudiantes expresen este resultado en el contexto del problema, indicando que las palabras tienen un componente emocional semejante o bien que no se puede rechazar esta hipótesis. Las respuestas encontradas son las siguientes:

Respuestas correctas:

RC.1.No podemos rechazar que las palabras tienen un componente emocional similar. Los estudiantes proporcionan una interpretación en el contexto del problema, tomando una decisión correcta, formulan esta expresión de diversas formas, por ejemplo:

“El componente emocional se distribuye homogéneamente en las diferentes palabras” (Estudiante 7).

“El componente emocional es igual en todas las palabras” (Estudiante 18).

“Todas las palabras afectan emocionalmente de la misma forma” (Estudiante 69).

Estos estudiantes han sido capaces de completar el último paso en el ciclo de modelización, que, según Henry (1997) consiste en interpretar los resultados del trabajo con el modelo. Por tanto traducen las conclusiones obtenidas mediante el trabajo matemático a conclusiones sobre aspectos de la realidad que se ha tratado de modelizar.

Respuestas parcialmente correctas:

RP.1. Los estudiantes no interpretan los resultados en el contexto del problema como se les pide, pero toman una decisión correcta haciendo uso de una interpretación genérica del contraste de hipótesis. Por tanto, aunque matemáticamente la solución es correcta, no llegan a completar el ciclo de modelización.

Es posible que ello sea un efecto del contrato didáctico, pues en las clases habituales de estadística el énfasis se pone en el cálculo estadístico y no tanto en la comprensión de qué implican estos cálculos sobre problemas de investigación de la especialidad del estudiante. Por ejemplo, observamos que el siguiente estudiante da una respuesta matemáticamente correcta, pero sin hacer referencia a la investigación planteada en el enunciado del problema:

“Los grados de libertad en este caso son 4. 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” (Estudiante 82).

Respuestas incorrectas:

RI.1. Los estudiantes interpretan en el contexto del problema como se les pide, pero toman una decisión incorrecta, rechazando la hipótesis nula y causando el *error tipo I* ya descrito anteriormente (véase la Sección 6.6.3), Un ejemplo se incluye a continuación, donde el estudiante ha sido capaz de traducir el resultado del trabajo matemático al contexto del problema, haciendo referencia al componente emocional de las palabras. Sin embargo, toma la decisión incorrecta pues rechaza la hipótesis nula, cuando el resultado no fue estadísticamente significativo:

“Según los datos, concluimos que los componentes emocionales de las tres palabras son distintos” (Estudiante 24).

RI.2. Los estudiantes no interpretan en el contexto del problema como se les pide en el enunciado del problema, y además toman una decisión incorrecta al rechazar la hipótesis nula y causando el *error tipo I* ya descrito anteriormente (véase la Sección 6.6.3). Una solución sería la que aparece a continuación, en que, como vemos, no hay interpretación en el contexto y además se rechaza la hipótesis nula:

“Rechazaríamos la hipótesis de independencia (porque este valor tan cercano a 0 indica un valor muy improbable de Chi-cuadrado)” (Estudiante 33).

RI.3. Los estudiantes interpretan en el contexto del problema, pero toman una decisión incorrecta. Además estos estudiantes para realizar sus conclusiones se basan en el valor p , pero no consideran el valor α explícitamente. Esta respuesta es parecida a la RI1; además, estos estudiantes no comparan estos dos valores para llegar a una toma de decisión, únicamente comparan la cercanía del valor p al valor 0, posiblemente el estudiante no recuerda la definición del nivel de significación o la regla de decisión en el contraste:

“Como el valor de $p=0,029$ es muy cercano a 0, indica un valor muy improbable de Chi-cuadrado se la H_0 es cierta, por tanto rechazamos la hipótesis de independencia” (Estudiante 15).

RI.4. Estos estudiantes no interpretan en el contexto del problema, y además no realizan una interpretación matemática de los resultados que obtienen, es decir, no indican si hay o no que rechazar la hipótesis. El estudiante no hace una toma de decisión, pues seguramente tiene un conflicto ya que han olvidado la lógica del proceso, por ejemplo:

“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 $p=0,029$ ” (Estudiante 95).

En la interpretación de los resultados según el contexto del problema, podemos observar en la Tabla 6.6.4.1 como 40 estudiantes (43,5%) realizan este apartado de forma correcta; además se puede apreciar en la tabla como esta frecuencia es la más alta. También resalta el hecho de que muchos estudiantes (25%) dejan el apartado sin contestar.

Tabla 6.6.4.1. Frecuencias y porcentajes de respuestas en el problema A, 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

Observamos que, en general, los estudiantes no llegan a interpretar correctamente el resultado en el contexto del problema, error ya señalado por diferentes autores (Birnbbaum, 1982; Vallecillos, 1994; Lecoutre, Lecoutre y Poitevineau, 2001; Haller y Kraus, 2002; Lecoutre, 2006). El error más frecuente encontrado fue que los estudiantes interpretan en el contexto del problema, pero toman una decisión incorrecta (RI.1), habiendo olvidado la lógica del contraste de hipótesis, en consecuencia, nos hace pensar que el estudiante tiene un conflicto cognitivo para llegar de la solución matemática a la solución particular del problema.

Síntesis de resultados en el problema A

Para resumir los resultados en este problema donde se pide a los estudiantes recorrer todos los pasos del contraste de hipótesis, observamos lo siguiente: 79,3% plantean hipótesis correctas o parcialmente correctas, en el primer apartado, donde podemos distinguir que 59,7% plantean correctamente la hipótesis nula y el 46,6% plantean correctamente la hipótesis alternativa, destacando al diferencia de porcentaje entre las respuestas correctas de ambas hipótesis.

63,1% determinan correctamente el estadístico y valor p (único punto facilitado por el software), en el apartado b). En este apartado si separamos las tareas a realizar, observamos que el estadístico Chi-cuadrado lo calcula correctamente el total de la muestra de estudiantes, los g.l. los calculan correctamente el 70,7% de la muestra y el p -valor el 63,1% de la muestra, lo que nos da una idea de la dificultad de las tres tareas. 51,9% toman la decisión correcta en el punto c y 43,5% interpretan correctamente los resultados en contexto, en el último apartado, 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 estudiantes 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 dicho 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 del componente emocional de las palabras) fue lo más difícil, de acuerdo a lo obtenido en otras investigaciones (por ejemplo, en Arteaga, 2011).

6.7. RESULTADOS EN EL PROBLEMA B: MEDIDAS DE ASOCIACIÓN EN TABLAS 2X2

Problema B. Se desea estudiar hasta qué punto existe relación entre el tiempo de residencia de inmigrantes en nuestro país y su percepción de integración. Se dispone de una muestra de 207 inmigrantes a los que se les evaluó en ambas variables obteniéndose la siguiente tabla de frecuencias observadas.

Tiempo de residencia	Grado de ansiedad		Total
	Bajo	Alto	
Menos tiempo	36	81	117
Más tiempo	81	9	90
Total	117	90	207

Calcule e interprete las medidas de asociación para tablas 2x2
(Adaptado de Guàrdia, Freixa, Però, y Turbany, 2007, pág. 130)

En este problema se pide a los estudiantes calcular e interpretar tres coeficientes de asociación para tablas 2x2, que estudiaremos de forma independiente. Para resolverlo, tienen a su disposición el programa de cálculo que, como recordamos, da los estadísticos calculados, una vez que el estudiante introduce la tabla de frecuencia. Por tanto, el estudiante sólo ha de reconocerlos e interpretarlos. A continuación analizamos detalladamente el tipo de respuestas obtenidas en el coeficiente Phi, y con menor detalle en el resto de los estadísticos pues las respuestas son similares en los diferentes estadísticos.

6.7.1. COEFICIENTE PHI DE PEARSON

El valor del coeficiente está cerca de -0,6, por lo cual la asociación es bastante alta (moderada-alta), y como el coeficiente es negativo la asociación es negativa. A continuación analizamos las respuestas obtenidas.

Respuestas correctas

RC.1. El estudiante calcula e interpreta bien el coeficiente, lo que implica una correcta implementación de las frecuencias absolutas en el programa y comprender cuál de los resultados ha de utilizar. Usa una notación adecuada (utilizando el símbolo Φ , signos, flechas...), y además deduce correctamente el signo del coeficiente, ya que recordemos que al ser el resultado de una raíz cuadrada el estudiante tiene que tomar la decisión del signo (Ato y López, 1996). Por tanto el estudiante tiene que combinar un procedimiento formal (la utilización de un coeficiente), con un procedimiento informal (como es la comparación de las celdas que informan de asociación directa e inversa). Interpreta correctamente, tanto el signo, como la intensidad. Un ejemplo es el siguiente:

“Phi de Pearson = -0,592 → La dependencia es inversa y moderada” (Estudiante 26).

Respuestas parcialmente correctas:

RP.1. Se interpreta adecuadamente el valor que obtiene, pero en el cálculo del coeficiente se confunde el signo, obteniendo el valor absoluto correcto. Aunque se usa una notación adecuada, se obtiene incorrectamente el signo del mismo, que tendría que haberse deducido mediante un procedimiento informal, y en este caso es negativo. El conflicto puede darse en los procedimientos informales ó en la utilización de raíces cuadradas, pues como sabemos por Euler (1770, p. 62)

“...la raíz cuadrada de cualquier número tiene siempre dos valores, uno positivo y el otro negativo; esto es que $\sqrt{4}$, por ejemplo, es igualmente 2 y -2, y en general, se puede adoptar tanto $-\sqrt{a}$ como $+\sqrt{a}$ para la raíz cuadrada de a ”.

Sin embargo, las investigaciones sobre la complejidad de interpretar el signo de las raíces cuadradas y su ambigüedad tienen una larga historia (Roach, Gibson y Weber, 2004). Un ejemplo se incluye a continuación:

“Coeficiente Phi de Pearson $\Phi = 0,592$. Dado que el coeficiente es positivo, la dependencia es directa y moderada” (Estudiante 22).

RP.2. El estudiante calcula mal el valor del coeficiente, además confunde el signo, pero realiza una interpretación correcta del valor que obtiene. En lugar de obtener el coeficiente Phi de Pearson del programa, lo calcula manualmente produciéndose un error en la fórmula de cálculo. Por otro lado, no es capaz de deducir el signo negativo del coeficiente mediante el uso de procedimientos informales, aunque interpreta correctamente el valor que obtiene. Se observan conflictos de cálculo de formulas, de signos de la raíz cuadrada y del uso de procedimientos informales. Un ejemplo es el siguiente:

“Phi de Pearson = Raíz de $(72,61)/207 = 0,04$; Observamos que es un valor positivo (dependencia directa). Es un valor moderadamente bajo” (Estudiante 102).

RP.3. Se calcula correctamente el coeficiente Phi de Pearson, obteniendo el signo y el valor correctos, pero se realiza una interpretación incorrecta, ya que no se reconoce los valores del coeficiente que corresponden a la asociación directa y cuáles de asociación inversa:

“ $\Phi = -0,592$. Como el valor es menor que 0, la dependencia es directa” (Estudiante 36).

RP.4. El estudiante es capaz de obtener el valor numérico correcto del coeficiente pero no deduce el signo de la raíz cuadrada, e interpreta mal el resultado. Aunque usa una notación adecuada, no reconoce los procedimientos informales que le informarían del signo del valor de asociación y además hace una interpretación de forma incorrecta del valor que obtiene. Se observan conflictos de cálculo del signo de la raíz cuadrada, del uso de procedimientos informales y de interpretación de los coeficientes de asociación. Un ejemplo lo vemos en el siguiente caso:

“Phi de Pearson = 0,592. Hay dependencia directa y perfecta” (Estudiante 19).

RP.5. Se usa una notación adecuada, pero no se realiza un procedimiento informal que le facilite el signo de la raíz cuadrada considerando el valor positivo, en lugar del negativo que es el correcto. Se observan conflictos de cálculo del signo de la raíz cuadrada, del uso de procedimientos informales y de interpretación de los coeficientes de asociación:

“Phi de Pearson = 0,592” (Estudiante 4).

RP.6. El estudiante calcula bien el valor numérico del coeficiente y deduce de forma correcta el signo de la raíz cuadrada. Usa una notación adecuada, y muestra el manejo de los procedimientos informales de las tablas de contingencia (los cuales le facilitan el signo negativo de este caso), Pero el estudiante no interpreta el valor del coeficiente, ya que no relaciona el resultado con el tipo de dependencia ó la intensidad de la misma. Un ejemplo sería:

“ $\Phi = -0,592$ ” (Estudiante 100).

Tabla 6.7.1.1. Frecuencias y porcentajes de respuestas en el problema B, apartado a

	Frecuencia	Porcentaje	Coeficiente Phi		
			% valor absoluto correcto	% signo correcto	% interpretación correcta
RC.1	33	35,8	35,8	35,8	35,8
RP.1	43	46,7	46,7		46,7
RP.2	3	3,3			3,3
RP.3	1	1,1	1,1	1,1	
RP.4	5	5,4	5,4		
RP.5	3	3,3	3,3		
RP.6	1	1,1	1,1	1,1	
No responde	3	3,3			
Total	92	100	93,4	38	85,8

Como resumen de los resultados en el cálculo e interpretación del coeficiente Phi de Pearson del problema B, podemos observar en la Tabla 6.7.1.1 que 33

estudiantes (35,8%) lo calculan e interpretan correctamente. Son mayoría los estudiantes que tienen respuesta parcialmente correcta (calculan bien el coeficiente y lo interpretan mal o al contrario). El cálculo del valor absoluto de Phi de Pearson y su interpretación resultaron sencillas (93,4% y 85,8% respectivamente). Por otro lado, la deducción del signo resultó difícil para los estudiantes, únicamente logrado por el 38%. Una respuesta muy frecuente resultó ser la respuesta (RP.1), en la que obtienen bien el valor numérico del coeficiente, sin deducir que el signo cambia, realizando una interpretación en este caso de forma correcta del resultado que obtiene.

6.7.2. RIESGO RELATIVO

Para calcular e interpretar el riesgo relativo, los estudiantes, con ayuda del programa calculan que toma el valor de 0,341, y coincide por filas y por columnas. Este hecho implica que es igual de probable predecir el tiempo de residencia si se sabe el grado de integración que al contrario. Lo cual también se ve por las celdas de la tabla, ya que son iguales las celdas b y c . Obtenemos categorías similares a las descritas en el apartado anterior, por lo que se describen muy resumidamente.

Respuestas correctas:

RC.1. El estudiante calcula e interpreta correctamente ambos valores, utilizando notación adecuada, e interpretando de forma correcta el significado de estos valores, es decir, indicando que cuando el coeficiente toma un valor menor que 1, existe una dependencia inversa entre las variables. La interpretación de este coeficiente también nos informa de cuanto es más probable la presencia de una variable respecto de la otra; si el alumno lo indica, muestra un uso correcto de nociones de probabilidad. Además, ha de manejar la complejidad de diferenciar las variables dependientes e independientes, proceso que es difícil, según Ruiz (2006), puesto que este coeficiente puede dar dos valores (por filas y por columnas). Un ejemplo se incluye a continuación, aunque el alumno no llega a realizar la interpretación completa, pero deduce la asociación negativa.

“El riesgo relativo nos indica que existe una asociación negativa. $0,34188034 < 1$ ” (Estudiante 9).

RC.2. Se calcula e interpreta correctamente los resultados obtenidos del riesgo relativo, e incluso se interpreta en el contexto del problema dado, algo poco frecuente en el resto de coeficientes. Además se utilizan competencias en probabilidad, ya que en la

interpretación de estos valores se requiere su uso. Un ejemplo sería:

“Riesgo relativo: RR columnas = 0,34188034, es 0,3418 veces más difícil tener menos tiempo de residencia con ansiedad baja. RR filas = 0,34188034, es 0,3418 veces más difícil tener más tiempo de residencia con ansiedad baja. En ambos casos tenemos que $RR < 1$, por lo tanto, existe una asociación negativa” (Estudiante 22).

Respuestas parcialmente correctas:

RP.1. El estudiante usa una notación adecuada y calcula correctamente los valores del riesgo relativo, pero no interpreta ninguno de los valores obtenidos, lo que podría deberse a algún conflicto en el uso de competencias probabilísticas:

“Riesgo Relativo = 0,3418034 Filas y columnas. No sabía acabar” (Estudiante 4).

RP.2. Aunque calcula correctamente el coeficiente, el estudiante los interpreta de forma incorrecta, como correspondiente a asociación. No ha sabido contrastar este resultado con las características ya comentadas de las frecuencias en la tabla, que claramente sugieren una asociación negativa:

“ $RR < 1$, 0,34 como es menor a 1, la asociación es positiva” (Estudiante 83).

R.P.3. El estudiante calcula de forma incorrecta los valores del riesgo relativo, pero interpreta bien el resultado que obtiene. El error se produce por cometer un error al introducir la tabla de contingencia en el programa (conflicto), lo que produce este error de cálculo mencionado, aunque usa una notación adecuada. La interpretación del valor que obtiene es correcta, pues deduce una asociación positiva, a pesar de no ser el valor correcto. No llega a especificar en forma probabilística el significado del valor del coeficiente.

“Los valores de riesgo relativo (RR) son por filas 6,92 y por columnas 2,92. Como las dos son mayores que 1 sabemos que existe asociación positiva” (Estudiante 91).

Tabla 6.7.2.1. Frecuencias y porcentajes de respuestas en el problema B, apartado b

	Frecuencia	Porcentaje	% RR correcto	% interpretación correcta
RC.1	60	65,2	62,7	62,7
RC.2	8	8,7	8,7	8,7
RP.1	3	3,3	3,3	
RP.2	5	5,4	5,4	
RP.3	7	7,6		7,6
No responde	9	9,8		
Total	92	100	79,9	78,8

En el cálculo, e interpretación, de los valores de los riesgos relativos (Ato y López, 1996), hemos podido observar como una gran parte de la muestra responden

correctamente. Deducen un valor correcto que interpretan adecuadamente 68 estudiantes (73,9%), siendo la frecuencia más alta de todas las categorías obtenidas. Además el 78,8% son capaces de interpretar el signo correctamente,

Sin embargo, casi ninguno hace una interpretación probabilística del significado del coeficiente en función de los riesgos de que aparezca o no un valor de la variable en función del otro. Este resultado posiblemente se deba a las dificultades de los estudiantes son las probabilidades condicionales, descritas, entre otros autores por Falk (1986), Díaz (2007) y Contreras (2011), pues los riesgos relativos son probabilidades condicionales y más aún, condicionales transpuestas, siendo la confusión entre una probabilidad condicional y su transpuesta uno de los errores frecuentes citados por los anteriores autores.

Exceptuando esta interpretación en términos de probabilidad condicional, se han tenido buenos resultados en la obtención del valor del riesgo relativo y su interpretación, pues aproximadamente el 80% responden correctamente estos dos puntos fundamentales en el proceso de resolución de este apartado.

6.7.3. RAZÓN DE PRODUCTOS CRUZADOS

El programa para el cálculo proporciona un valor de 0,0493 para este coeficiente. Se ha de interpretar este resultado y llegar a la conclusión de que, como es inferior al valor 1, hay una asociación inversa entre las variables, lo que indica que la razón entre el número de emigrantes con “menos tiempo” de residencia y “más tiempo” de residencia es menor cuando el grado de ansiedad es bajo. A continuación se comentan muy brevemente las categorías.

Respuestas correctas:

RC.1. El estudiante calcula e interpreta de forma correcta el signo del coeficiente de la razón de productos cruzados, por lo que utiliza razones que son necesarias en dicha interpretación. Como en el caso anterior, no llega a dar una interpretación en el contexto del problema, lo que será constante en el resto de las categorías.

“Razón de productos cruzados: 0,04938272. Como $RC < 1$, esto nos indica que hay dependencia inversa” (Estudiante 14).

Respuestas parcialmente correctas:

RP.1. El estudiante calcula bien el valor de la razón de productos cruzados, pero hay un problema de interpretación que no realiza, lo que sugiere algún conflicto sobre la

comprensión del coeficiente ó en la noción de razón. Un ejemplo se incluye a continuación:

“Razón de productos cruzados = 0,04938272” (Estudiante 8).

RP.2. En este caso, aunque el cálculo del valor de la razón de productos cruzados es correcto, se realiza una interpretación incorrecta, indicando que la asociación es directa, no habiendo comprendido el significado del coeficiente. Un ejemplo sería:

“Razón de productos cruzados = 0,04938272. Existe asociación directa” (Estudiante 19).

RP.3. El estudiante calcula de forma incorrecta la razón de productos cruzados, por lo que tiene algún conflicto en la construcción de la tabla de contingencia en el programa ó en la aplicación de la fórmula (como ocurre en el ejemplo que se intercambia el numerador y el denominador al operar), pero la interpretación del valor que obtiene es correcta. Un ejemplo sería el siguiente, que, además, es uno de los pocos estudiantes que da una interpretación probabilística del coeficiente, mostrando una buena comprensión de la probabilidad condicional y la diferenciación de la variable dependiente e independiente.

“Razón de productos cruzados es 20,25, y esto implica que la razón entre los casos que aparecen A y no A es mayor cuando está presente B” (Estudiante 86).

Tabla 6.7.3.1. Frecuencias y porcentajes de respuestas en el problema B, apartado c

	Frecuencia	Porcentaje	% cálculo correcto RR	% interpretación correcta
RC.1	54	58,7	58.7	58.7
RP.1	10	10,9	10.9	
RP.2	7	7,6	7.6	
RP.3	6	6,5		6.5
No responde	15	16,3		
Total	92	100	77.2	65.2

En el cálculo, e interpretación, de la razón de productos cruzados (Ato y López, 1996), hemos podido observar que ninguno de los estudiantes dio una respuesta totalmente incorrecta, obteniendo siempre bien una de las dos partes principales: el valor (77,2%) ó realizando una interpretación correcta del signo (65,2%), incluso en casos donde el valor obtenido fue erróneo. Pero son pocos los que dan una interpretación en el contexto del problema o bien una interpretación en términos de probabilidad.

Síntesis de resultados en el problema B

Los resultados en el cálculo e interpretación del signo de las medidas de asociación para tablas 2x2 han sido bastante buenos, puesto que el 83,7% de estudiantes intentan calcular las tres medidas de asociación para tablas 2x2; el 96,7% el coeficiente Phi de Pearson y el 90,2% el Riesgo relativo.

Al comparar los resultados en los tres coeficientes, observamos que el coeficiente Phi es calculado y su signo interpretado correctamente por el 35,8 y parcialmente por el 60,4; el riesgo relativo correctamente por el 73,9 y parcial el 16,3 y la razón de productos cruzados correctamente por el 58,7 y parcial por el 25%.

En resumen, el signo e intensidad de estas medidas de asociación parecen sencillos de comprender, aplicar e interpretar, siendo la más fácil el riesgo relativo, seguido por la razón de productos cruzados y finalmente el coeficiente Phi. Los estudiantes fueron capaces de calcular la intensidad de la asociación en una tabla de contingencia, utilizando procedimientos y cálculos matemáticos. Para ello trabajaron con objetos matemáticos complejos en los que se basan estos coeficientes.

Finalmente resaltamos que en este problema son pocos los estudiantes que llegan a interpretar los resultados en el contexto. Salvo raras excepciones, la interpretación se limita a indicar la intensidad y signo, sin dar una interpretación probabilística, posiblemente por dificultades con el razonamiento sobre probabilidad condicional. En el único caso que se obtuvieron algunas interpretaciones de este tipo fue para los valores de los riesgos relativos. Sin embargo, pensamos que la interpretación de los modelos matemáticos es difícil, según lo señalado por Henry (1997).

6.8. RESULTADOS EN EL PROBLEMA C: MEDIDAS DE ASOCIACIÓN EN TABLAS RXC

Problema C. Se está realizando un estudio sobre la población de estudiantes que hicieron la prueba de selectividad en el curso 1978-1979. Dos de las variables registradas fueron “convocatoria” (junio-septiembre) y “tipo de estudios” por el que se inclinaba el estudiante (grado medio, facultades, escuelas técnicas). Los resultados encontrados en la Universidad Complutense fueron los expuestos en la tabla siguiente:

	Escuelas Técnicas	Facultades	Grado Medio
Septiembre	200	500	800
Junio	1500	1300	700

Calcule e interprete las medidas de asociación para tablas rxc
(Adaptado de San Martín y Pardo, 1989, pág. 449)

En este problema se pide a los estudiantes calcular e interpretar tres coeficientes de asociación para tablas rxc, que estudiaremos de forma independiente. Para resolverlo, tienen a su disposición el programa de cálculo, por lo que la principal dificultad es la aplicación e interpretación. Al ser las categorías muy similares a las del problema B se comentan muy resumidamente

6.8.1. COEFICIENTE DE CONTINGENCIA DE PEARSON

Con ayuda del programa, se calcula el valor del coeficiente que es igual a 0,342. Además el estudiante tiene que calcular el valor máximo, que también lo da el programa siendo igual 0,7071. El estudiante tiene que concluir que el valor observado del coeficiente está un poco por debajo de la mitad de su posible máximo; por tanto podemos considerar que hay una dependencia moderada ó moderada-baja.

Respuestas correctas

RC.1. Se calcula e interpreta y de forma correcta el valor del coeficiente de contingencia de Pearson y su valor máximo, que es necesario conocer para poder interpretar adecuadamente la intensidad de la asociación. (Ruiz-Maya, Martín, Montero y Uriz, 1995), Para ello, se introduce correctamente los datos en el programa, así como el número de filas y de columnas. Un ejemplo es el siguiente, en que el alumno deduce correctamente la existencia de una asociación moderada.

“Coeficiente de contingencia de Pearson = 0,342. Está asociación es moderada porque el máximo es 0,70710078 y sólo llegamos a la mitad” (Estudiante 8).

Respuestas parcialmente correctas

RP.1. El estudiante calcula bien los valores necesarios, es decir el del coeficiente y su posible máximo, pero no interpreta la intensidad de la asociación. Un ejemplo, en que el alumno correctamente deduce que hay dependencia al obtener un coeficiente distinto de cero, pero no es capaz de decidir qué intensidad indica el valor hallado, se incluye a continuación. El alumno explícitamente indica el rango posible de valores que, para la tabla dada, puede tomar el coeficiente.

“Coeficiente de contingencia de Pearson = 0,342. Existe dependencia porque difiere de 0. Max (C) = 0,70710678. Rango entre (0, 0’7) ” (Estudiante 15).

RP.2. Este es un caso particular del anterior. El estudiante calcula bien el valor del coeficiente de contingencia de Pearson y el máximo que este valor puede alcanzar,

pero surge un problema, pues la interpretación del valor obtenido es incorrecta, ya que indica que la asociación es moderada-alta, en lugar de una asociación moderada ó moderada-baja. Un problema en esta interpretación es que no hay una regla fija que el estudiante pueda aplicar (por ejemplo, usar unos valores fijos a partir de los cuáles la asociación sea alta o moderada), pues la decisión sobre la intensidad que debe considerarse alta es hasta cierto punto subjetiva. No obstante, tradicionalmente sólo los valores mayores que 0,7 se consideran altos, por lo que es posible que el estudiante haya extrapolado indebidamente esta propiedad (aplicable al coeficiente) a su valor máximo, que si es superior a 0,7.

“Coeficiente de contingencia de Pearson = 0,342, Máximo = 0,7071. Se obtienen valores estadísticamente significativos, por lo que la asociación es moderada-alta” (Estudiante 53).

RP.3. El estudiante calcula bien el coeficiente de contingencia de Pearson, aunque comete un error en el valor máximo que puede tener este coeficiente, posiblemente debido a confundir el número de columnas y de filas o no introducirlo correctamente en el programa. Estos estudiantes dan una interpretación correcta de los valores que han obtenido, considerando una asociación moderada ó moderada-baja, ya que el coeficiente no supera la mitad del valor máximo.

“Coeficiente de contingencia de Pearson = 0,3424, Max = 0,81649658. Valor moderadamente bajo por tanto la asociación es moderadamente baja” (Estudiante 65).

RP.4. El estudiante calcula correctamente el coeficiente de contingencia de Pearson, pero comete un error en el cálculo de su valor máximo, y además realiza una interpretación errónea de su magnitud, considerando una asociación moderada-alta, en lugar de una asociación moderada ó moderada-baja. Un ejemplo sería el siguiente, en el que también podría darse el conflicto anteriormente citado de extender indebidamente una propiedad del coeficiente a su máximo.

“Coeficiente de contingencia de Pearson = 0,342, Máximo = 0,81649658. La asociación es moderada-alta” (Estudiante 76).

RP.5. Al igual que en los dos casos anteriores, el estudiante calcula bien el valor del coeficiente de contingencia de Pearson, y confunde el valor máximo y en este caso no interpreta. Un ejemplo lo mostramos a continuación:

“C = 0,342, Max = 0,8165” (Estudiante 36).

RP.6. Se calcula bien el coeficiente de contingencia de Pearson, lo que implica una buena construcción de la tabla, pero no el máximo, y la interpretación es incorrecta

ó no la hace. Los conflictos son: no considerar el número de filas y columnas (ya que el valor máximo es incorrecto), y el significado del valor obtenido (producido por no interpretar ó hacerlo de forma errónea). Un ejemplo se incluye a continuación.

“C. contingencia de Pearson = 0,342. Podemos decir que existe una asociación moderada-alta” (Estudiante 19).

RP.7. El estudiante calcula mal el coeficiente de contingencia de Pearson (lo que implica un error al introducir la tabla en el ordenador), pero calcula bien el máximo posible de este coeficiente. No interpreta los valores que obtiene (con lo que hay algún conflicto en esta interpretación). Una respuesta de este tipo sería:

“Coeficiente de contingencia de Pearson, $C = \text{Raíz cuadrada de } (\chi^2 / (\chi^2 + n)) = 0,284$. Máximo: $\text{Max}(C) = \text{Raíz cuadrada de } (\text{Min}(r-1, c-1) / (1 + \text{Min}(r-1, c-1))) = 0,70710678$ ” (Estudiante 16).

RP.8. Aquí el estudiante comete un error de cálculo del coeficiente como anteriormente y calcula bien su máximo como el caso anterior, pero hace una interpretación correcta de los valores que ha obtenido. Una respuesta en esta categoría se muestra a continuación:

“Coeficiente de contingencia de Pearson = 0,374. $\text{Max}(C) = 0,707$. Asociación moderada entre ambas variables” (Estudiante 7).

Respuestas incorrectas

RI.1. Calcula o interpreta incorrectamente los dos valores. Esta respuesta sugiere un conflicto en la construcción de la tabla de contingencia al introducirla al ordenador y respecto al valor máximo que este coeficiente puede alcanzar cuando existe una asociación perfecta, el error sugiere que no considera el número correcto de columnas y número de filas;

“Coeficiente de contingencia de Pearson = 0,196, $\text{Max} = 0,81$. Viendo el máximo resultante, 0,81, el nivel de asociación es moderada-alta” (Estudiante 88).

RI.2. Calcula los coeficientes correspondientes a la asociación de tablas 2x2. El estudiante calcula los coeficientes de asociación de tablas 2x2 (coeficiente Phi de Pearson, riesgo relativo y razón de productos cruzados), en lugar de los adecuados para tablas rxc (coeficiente de contingencia de Pearson, V de Cramer y Lambda de Goodman y Kruskal). Estos estudiantes tienen un conflicto al no saber cuándo aplicar los coeficientes de asociación para tablas 2x2, y para tablas rxc.

Tabla 6.8.1.1. Frecuencias y porcentajes de respuestas en el coeficiente de contingencia de Pearson

	Frecuencia	Porcentaje	% cálculo coeficiente correcto	% cálculo Max correcto	% interpretación correcta
RC.1	36	39,1	39,1	39,1	39,1
RP.1	2	2,2	2,2	2,2	
RP.2	13	14,1	14,1	14,1	
RP.3	7	7,6	7,6		7,6
RP.4	5	5,4	5,4		
RP.5	1	1,1	1,1		
RP.6	4	4,3	4,3		
RP.7	1	1,1		1,1	
RP.8	2	2,2		2,2	2,2
RI.1	4	4,3			
RI.2	2	2,2			
No responde	15	16,3			
Total	92	100	74,8	58,7	48,9

Podemos observar en la Tabla 6.8.1.1 que sólo 36 estudiantes (39,1%) responden correctamente al cálculo e interpretación del coeficiente de contingencia de Pearson. La respuesta correcta es la más frecuente, seguida de RP.2, en la que se calculan bien los valores (del coeficiente y su máximo posible), pero la interpretación de estos resultados obtenidos es incorrecta.

También hacemos notar que el coeficiente de contingencia de Pearson ha sido calculado correctamente por el 74,8% de los estudiantes, su máximo es realizado correctamente por el 58,7% y la interpretación la realiza correctamente el 48,9% de los estudiantes. Como en problemas anteriores, observamos que la mayor dificultad se da en la interpretación.

6.8.2. V DE CRAMER

Para calcular e interpretar el valor V de Cramer, los estudiantes con ayuda del programa calculan el valor 0,364, y posteriormente han de interpretar que está un poco por debajo de la mitad de su valor máximo (el valor 1). Por ello, de este coeficiente se deduce que hay una dependencia moderada ó moderada-baja. A continuación comentamos muy brevemente las respuestas obtenidas.

Respuestas correctas:

RC.1. El estudiante interpreta y calcula el valor del coeficiente (V de Cramer = 0,364) correctamente. El estudiante usa una notación adecuada, construye de forma adecuada la tabla de contingencia en el programa, tiene en consideración el número de filas y de columnas, y además considera correctamente su máximo como uno (Ruiz-Maya, Martín, Montero y Uriz, 1995). Posteriormente, el estudiante realiza una

interpretación adecuada del resultado. Un ejemplo se incluye a continuación.

“V de Cramer, $V = 0,364$. Dependencia moderada” (Estudiante 29).

Respuestas parcialmente correctas:

RP.1. En este caso se calcula bien el coeficiente V de Cramer, pero la interpretación es errónea. En el siguiente ejemplo hay un conflicto de interpretación ya que se está considerando independencia, en lugar de una asociación moderada ó moderada-baja:

“V de Cramer = 0,364, independencia” (Estudiante 33).

RP.2. Esta respuesta es un caso particular de la anterior, pues el estudiante, calcula bien el valor del coeficiente de V de Cramer, pero no interpreta el resultado, como vemos en el siguiente ejemplo:

“V de Cramer = 0,364” (Estudiante 86).

RP.3. El estudiante calcula mal el valor del coeficiente; tiene un conflicto o su máximo, pero da una interpretación correcta de los valores que han obtenido. Como en el ejemplo que sigue, un resultado muy frecuente fue el 0,515, que se obtiene al considerar 3 filas y 3 columnas, en lugar de 2 filas y 3 columnas. La interpretación, sin embargo es adecuada.

“V de Cramer = 0,515. Asociación moderada entre ambas variables” (Estudiante 7).

Respuestas incorrectas:

RI.1. Un estudiante, calcula incorrectamente el valor del coeficiente debido a un conflicto en la construcción de la tabla. Además no realizan ningún tipo de interpretación del resultado que obtiene:

“V de Cramer = 0,838” (Estudiante 82).

RI.2. Este sería un caso particular del anterior, pues se calcula mal el coeficiente (como anteriormente) y la interpretación que hace es errónea. Un ejemplo se incluye a continuación.

“V de Cramer = 0,419. Independencia” (Estudiante 4).

RI.3. Calcula e interpreta las medidas de asociación para tablas 2x2. El estudiante no distingue los coeficientes que tiene que usar para tablas 2x2, y los coeficientes que tiene que usar para las tablas rxc.

Tabla 6.8.2.1. Frecuencias y porcentajes de respuestas en el coeficiente V de Cramer

	Frecuencia	Porcentaje	% cálculo de V de Cramer correcto	% interpretación de V de Cramer correcta
RC.1	41	44,5	44,5	44,5
RP.1	14	15,2	15,2	
RP.2	2	2,2	2,2	
RP.3	9	9,8		9,8
RI.1	1	1,1		
RI.2	6	6,5		
RI.3	1	1,1		
No responde	18	19,6		
Total	92	100	61,9	54,3

Podemos observar, en la Tabla 6.8.2.1, como únicamente 41 estudiantes (44,5%) calculan e interpretan correctamente del valor V de Cramer, a pesar de disponer del programa de Excel. Entre las respuestas incorrectas destaca la respuesta RP.1, en la que los estudiantes calculan bien el coeficiente de la V de Cramer, pero la interpretación que realizan es errónea. Una de las problemáticas se produce al introducir la función $\text{Mín}\{\}$ ó al introducir el número de filas y columnas en el programa de ordenador. El problema fue de dificultad media, pues la obtención de la V de Cramer la realiza correctamente el 61,9% de estudiantes y su interpretación (54,3%).

6.8.3. LAMBDA DE GOODMAN Y KRUSKAL

Con ayuda del software se calcula el coeficiente Lambda de Goodman y Kruskal, que toma el valor 0,067. El estudiante tiene que interpretar que el error que se ve reducido al predecir el valor de la variable dependiente Y conocido el valor de la variable independiente X , es 6,7%.

Respuestas correctas:

RC.1. El estudiante calcula e interpreta correctamente el valor de Lambda de Goodman y Kruskal, como proporción del error reducido. Usa una notación adecuada, e introduce la tabla de contingencia de forma correcta en el programa, informando en la interpretación que es el porcentaje de error que se ve reducido al predecir el valor de la variable dependiente Y , conocido el valor de la variable independiente X en nuestro caso el 6,7% (Ruiz-Maya, Martín, Montero y Uriz, 1995). Un ejemplo sería:

“Lambda de Goodman y Kruskal = 0,067. Hemos reducido un 6,7% con respecto al error que teníamos” (Estudiante 55).

RC.2. Otra posible solución correcta, aunque más pobre que la anterior es que el estudiante calcule correctamente el valor de Lambda de Goodman y Kruskal, y lo interprete en función de si hay asociación o no. Una solución de este tipo es la siguiente en que el estudiante interpreta el coeficiente únicamente como coeficiente de asociación y no lo relaciona con la reducción en el porcentaje de error en la predicción:

“Lambda de Goodman y Kruskal = 0,067. La asociación es bastante baja” (Estudiante 39).

Respuestas parcialmente correctas:

RP.1. En este caso, el estudiante calcula bien el valor del coeficiente Lambda de Goodman y Kruskal, pero no realizan una interpretación de este resultado, por lo que pensamos que no ha comprendido la forma de interpretarlo. A continuación vemos un caso de este tipo:

“Lambda de Goodman y Kruskal = 0,067” (Estudiante 31).

RP.2. Como en el caso anterior, el estudiante calcula bien el valor de Lambda de Goodman y Kruskal, pero hace interpretación errónea, en términos de asociación, considerando una asociación moderada, ó bien en términos de predicción de error, indicando que no mejora la predicción con el conocimiento de una variable. Un ejemplo se incluye a continuación.

“Lambda de Goodman y Kruskal = 0,067. La asociación es moderada” (Estudiante 35).

Respuestas incorrectas:

RI.1. El estudiante realiza un cálculo incorrecto del valor de Lambda de Goodman y Kruskal, lo cual viene motivado por una construcción de la tabla de forma errónea. Además no interpreta el coeficiente, sugiriendo un desconocimiento de su interpretación. Un ejemplo sería el que reproducimos a continuación:

“Lambda de Goodman y Kruskal = 0,500” (Estudiante 82).

RI.2. Calcula e interpreta las medidas de asociación para tablas 2x2. El estudiante no distingue los coeficientes que tiene que usar para tablas 2x2, y los coeficientes que tiene que usar para las tablas rxc.

RI.3. No calcula Lambda de Goodman y Kruskal, debido a que consideraron que no es una medida de asociación, pues en los apuntes se presentaba como una medida basada en la reducción proporcional del error.

Tabla 6.8.3.1. Frecuencias y porcentajes de respuestas en el coeficiente Lambda

	Frecuencia	Porcentaje	% cálculo de Lambda correcto	% interpretación correcta de Lambda
RC.1	29	31,6	31,6	31,6
RC.2	15	16,3	16,3	16,3
RP.1	11	11,9	11,9	
RP.2	11	11,9	11,9	
RI.1	2	2,2		
RI.2	1	1,1		
RI.3	5	5,4		
No responde	18	19,6		
Total	92	100	71,7	47,9

En la Tabla 6.8.3.1 presentamos los resultados, viendo como 44 estudiantes (47,9%) dan la respuesta correcta. Como se ha visto, algunos estudiantes calcularon únicamente el coeficiente de contingencia de Pearson y el valor V de Cramer, ya que el coeficiente Lambda es una medida basada en la reducción proporcional del error, y no la consideraron como coeficiente de asociación. Entre las respuestas más frecuentes destacan, principalmente, las correctas y, posteriormente, las parcialmente correctas, lo cual muestra unos resultados buenos.

El cálculo correcto de Lambda ha sido bastante alto (71,7% de la muestra), pero como en otros coeficientes la interpretación correcta de Lambda sigue siendo baja (47,9% de la muestra).

Síntesis de resultados en el problema C

En la Sección 6.7 hemos visto que una mayoría de estudiantes reconocían las medidas de asociación para tablas 2x2, excepto el Riesgo Relativo, sin embargo, cuando observamos las respuestas a las medidas de asociación para tablas rxc, vemos que el porcentaje de no respuestas es cercano al 20%, lo que nos indica que los estudiantes reconocen mejor las medidas de asociación para tablas 2x2 que las medidas de asociación para tablas rxc.

Al comparar los tres coeficientes observamos que el coeficiente de contingencia de Pearson es calculado e interpretado correctamente por el 39,1%, parcialmente por el 38,1% y de forma incorrecta por el 6,5%; el valor V de Cramer correctamente por el 44,5%, parcialmente por el 27,2% y de forma incorrecta por el 8,7%; y el valor Lambda de Goodman y Kruskal correctamente por el 47,9%, parcialmente por el 29,2%, y de forma incorrecta por el 3,3%.

En resumen estas medidas de asociación parecen de dificultad moderada, siendo la más sencilla la Lambda de Goodman y Kruskal, seguida por la V de Cramer y finalmente el coeficiente de contingencia de Pearson. Posiblemente la mayor facilidad de calcular e interpretar el valor de Lambda se debe a que este valor no se veía afectado por el número de filas y columnas, y no requiere el cálculo de un valor máximo de forma independiente. Un problema que se repite respecto a los anteriores ítems es la dificultad de interpretación de los modelos matemáticos, señalada por Henry (1997).

6.9. RESULTADOS EN EL PROBLEMA D: CÁLCULO DE PROBABILIDADES Y JUICIOS DE ASOCIACIÓN

Problema D. Un grupo de 200 personas aquejadas de ansiedad fue dividido aleatoriamente en dos subgrupos. Al primer grupo se ofreció unas píldoras realmente efectivas para que tomaran tres al día, y al otro se ofreció un placebo (medicamento sin efecto). Al cabo de un mes fueron interrogados sobre la eficacia de las pastillas tomadas, con el siguiente resultado:

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

- Preguntado un paciente, dice que su ansiedad casi ha desaparecido; está encantado con el tratamiento. ¿Cuál es la probabilidad de que este paciente haya tomado el placebo?
- ¿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, que se realiza sin ayuda del programa de ordenador encontramos tres apartados los cuales estudiaremos de forma independiente. En el primero, se plantea al estudiante el cálculo de una probabilidad condicional a partir de una tabla de contingencia, problema abordado también por Contreras (2011), quien encontró numerosos conflictos semióticos en su resolución y compararemos nuestros resultados con los suyos. En el resto de apartados se pide un juicio de asociación y su justificación, lo que permitirá comparar con los resultados obtenidos en los juicios de asociación realizados antes de la enseñanza (Capítulo 3).

Para resolver el primer apartado el estudiante ha de leer e interpretar los valores de la tabla, la solución correcta sería: $P(\text{Placebo} | \text{ansiedad ha disminuido}) = 0,657$. Para realizar el juicio de asociación, en el segundo apartado podría llevar a cabo un contraste de independencia formal, calculando el estadístico Chi-cuadrado, o bien simplemente comparar las frecuencias observadas con las esperadas en caso de independencia; analizar los valores de las celdas que informan de una dependencia directa (50 y 39), y compararlos con los valores de las celdas que informan de una dependencia inversa (15

y 96). También podrían utilizarse cualquiera de las estrategias correctas que se observaron en el estudio inicial (Capítulo 3), es decir las de nivel 5 (Pérez Echeverría, 1990) en que se usen todas las celdas de la tabla con comparaciones multiplicativas.

6.9.1. CÁLCULO DE PROBABILIDAD

Como se ha indicado el estudiante ha de leer la tabla para obtener la probabilidad pedida. A continuación se describen las respuestas encontradas.

Respuestas correctas:

RC.1. El estudiante usa una notación adecuada, calculando de forma directa la probabilidad condicionada a partir de una tabla de contingencia. El estudiante identifica la pregunta como de probabilidad condicionada, mostrando que diferencia este concepto de la probabilidad simple y conjunta. También ha sido capaz de leer la tabla correctamente, mostrando un nivel de lectura adecuado (lectura entre los datos, según Curcio, 1989); identifica el numerador y denominador de la regla de Laplace, es decir, los casos favorables y posibles a partir de la tabla de contingencia. Es por tanto capaz de localizar la intersección de dos sucesos, su frecuencia conjunta y la frecuencia marginal. Un ejemplo sería:

$$“P(\text{Placebo/ansiedad ha disminuido}) = \frac{96}{(50 + 96)} = \frac{96}{146} = 0,657” \text{ (Estudiante 62).}$$

RC.2. El estudiante usa una notación adecuada, calculando la probabilidad a partir de la fórmula de la probabilidad condicionada que recuerda de las asignaturas pasadas. Como en el caso anterior, muestra que es capaz de discriminar probabilidades simples, compuestas y condicionales y leer la tabla a nivel adecuado según Curcio (1989). Además, es capaz de recordar y aplicar correctamente la fórmula de la probabilidad condicional, calculando primero la probabilidad conjunta del numerador, y después la probabilidad simple del denominador, en cada una de las cuáles ha aplicado correctamente la regla de Laplace. A continuación mostramos una respuesta de este tipo:

$$“P(\text{Placebo/ansiedad ha disminuido}) = \frac{P(\text{Placebo} \cap \text{ansiedad}_\text{menor})}{P(\text{ansiedad}_\text{menor})} = \frac{0,48}{0,73} = 0,657” \text{ (Estudiante 21).}$$

Respuestas parcialmente correctas:

RP.1. El estudiante identifica correctamente que se trata de calcular una probabilidad condicional, pero comete un error de cálculo. Contreras (2011) encontró 3 casos (1,6%) que cometen este mismo error, considerando como “básicamente correctas” el caso que se plantea bien la fórmula pero tiene errores de cálculo, este autor. Mostramos una respuesta en esta categoría en lo que sigue:

“0,675” (Estudiante 2).

Respuestas incorrectas:

RI.1. El estudiante calcula la probabilidad conjunta, en lugar de la probabilidad condicional, confundiendo estos dos conceptos. Pollatsek, Well, Konold y Hardiman (1987) consideran que este error surge por la falta de precisión del lenguaje verbal de los enunciados de los problemas cuando expresamos probabilidades condicionales. Asimismo Einhorn y Hogarth (1986), observaron que algunos estudiantes, confundían la interpretación de estas dos probabilidades, y la mitad de los sujetos del estudio de Ojeda (1995) interpretaron la intersección como condicionamiento. También aparece en el estudio de Contreras (2011) donde se dan 14 casos (7,7%), bastantes menos que en nuestro estudio. Una respuesta sería la siguiente:

“ $\frac{96}{200} = 0,48$ ” (Estudiante 19).

RI.2. El estudiante calcula la probabilidad simple, en lugar de la probabilidad condicional, mostrando confusión entre estos dos conceptos. Como hemos indicado Pollatsek, Well, Konold y Hardiman (1987), consideran que muchos errores de cálculo de la probabilidad condicional surgen por la confusión entre probabilidad condicional y simple o conjunta. Este conflicto aparece también en otras investigaciones (Estrada y Díaz, 2007; Contreras, 2011); en este último trabajo sólo aparece en 2 sujetos (1,1%). Un ejemplo es el que reproducimos a continuación:

“ $\frac{146}{200} = 0,78$ ó $\frac{135}{200} = 0,67$ ” (Estudiante 14).

RI.3. El estudiante calcula una probabilidad condicional diferente de la que se le pide. Batanero, Estepa, Godino y Green (1996), observaron la confusión entre las dos frecuencias condicionales relacionadas con una misma celda de la tabla de contingencia, lo que en nuestro caso podría producir este error. Como en otras respuestas anteriores Falk (1986) considera como explicación de este error, que denomina *falacia de la*

condicional transpuesta la imprecisión del lenguaje cotidiano. Este error sigue apareciendo en más investigaciones, como es el caso de Contreras (2011), donde 9 estudiantes (4,9%) confunden $P(A|B)$ con $P(B|A)$. En nuestro estudio se produce cuando el estudiante presenta un conflicto de aplicación de la fórmula de la probabilidad condicional, ya que intercambia los términos. Un caso sería:

$$\text{“} \frac{50}{146} = 0,342 \text{” ó “} \frac{96}{135} \text{” (Estudiante 39).}$$

RI.4. El estudiante divide dos frecuencias absolutas, que no llevan a una probabilidad sino a una razón; es decir, confunde estos dos conceptos. En la investigación de Contreras (2011), se encuentra un sujeto (0,5%) con una confusión de este tipo, es decir calcular $P(A/B) = \frac{N(A \cap B)}{N(\overline{A \cap B})}$. El conflicto se produce al no recordar

la fórmula de la probabilidad condicional. A continuación se muestra un ejemplo:

$$\text{“} \frac{50}{96} = 0,52 \text{” (Estudiante 54).}$$

RI.5. El estudiante confunde la fórmula de la probabilidad condicional, intercambiando el numerador con el denominador. Esto le lleva a obtener una probabilidad superior a uno, por lo que no es consciente de los axiomas de la probabilidad. Contreras (2011) también describe 2 casos (1,1%) de estudiantes que obtienen probabilidades mayores que uno, pero ninguno que intercambie numerador y denominador en la fórmula de la probabilidad condicional. El estudiante tiene un conflicto en no recordar la fórmula de la probabilidad condicional, y otro en considerar que una probabilidad puede superar la unidad, dando respuestas como la siguiente:

$$\text{“} \frac{146}{96} = 1,52 \text{” (Estudiante 100).}$$

Tabla 6.9.1.1. Frecuencias y porcentajes de respuestas en el problema 4, apartado a

	Frecuencia	Porcentaje
RC.1	55	58,5
RC.2	1	1,1
RP.1	1	1,1
RI.1	26	27,6
RI.2	5	5,3
RI.3	2	2,1
RI.4	3	3,2
RI.5	1	1,1
Total	94	100

Como resumen del cálculo de la probabilidad condicionada, en la Tabla 6.9.1.1 observamos 56 (59,6%) respuestas correctas. Entre las incorrectas resalta la respuesta RI.1, en la que los estudiantes calculan la probabilidad conjunta, en lugar de la probabilidad condicional, error que, como se ha dicho, aparece en bastantes investigaciones. El resto de errores apenas aparece. También queremos destacar que este apartado ha sido respondido por todos los estudiantes, lo que nos indica que, en principio, la tarea de calcular una probabilidad condicionada fue asequible.

En la investigación de Contreras (2011) se obtiene un 43,7% de resultados correctos en futuros profesores de primaria, y Estrada y Díaz (2007) obtuvieron un 56% de respuestas correctas en el cálculo de probabilidades condicionales en tablas de contingencia también con futuros profesores. Por otro lado, Díaz (2007), obtiene un 52% de respuestas correctas en estudiantes de psicología. Nuestros resultados (59,6% de respuestas correctas), son mejores que todas estas investigaciones, lo que sugiere que la enseñanza fue productiva en nuestros estudiantes.

6.9.2. JUICIOS DE ASOCIACIÓN.

Solo hemos considerado dos tipos de respuestas en este apartado. En este problema, se presenta una pequeña asociación (Φ de Pearson=0,061, $RR_{columnas}=1,2329$, $RR_{filas}=1,0817$ y razón de productos cruzados=1,3541), por lo que algunos estudiantes consideran que no la hay, aunque posteriormente se puede apreciar como los estudiantes en su mayoría al aplicar propiedades de independencia y observar que no se cumplen con exactitud, tienden a indicar la asociación.

Tabla 6.9.2.1. Frecuencias y porcentajes de respuestas en el problema D, apartado b

	Frecuencia	Porcentaje
RC.1 Asociación	68	72,3
RI.1 Independencia	12	12,8
No responde	14	14,9
Total	94	100

Los resultados sobre el juicio de asociación se incluyen en la Tabla 6.9.2.1, donde se observa que la mayoría de los estudiantes indica que hay asociación, pues la independencia no es perfecta. Ello nos indica un buen aprendizaje de las nociones de asociación, independencia y sus propiedades, pues la mayoría detecta la asociación, a pesar de ser pequeña.

Al comparar con los ítems de tablas 2x2 del cuestionario de conocimientos intuitivos (Capítulo 3), vemos como en aquel caso teníamos el caso de independencia (donde el 78,1% consideraban relación y el 21,7% consideraban independencia), el caso de dependencia inversa (donde el 96,1% consideraban relación y el 3,4% consideraban independencia), y el caso de dependencia directa (donde el 93,2% consideran relación y el 5,8% consideran independencia). Aparentemente los resultados son similares, pero la intensidad de la asociación no es similar a la de ninguno de los ítems de dicho estudio. Además en aquel cuestionario no se diferenciaba en el signo de la asociación, haciendo mucho más difícil el problema que tratamos en este capítulo.

6.9.3. ESTRATEGIA EN EL JUICIO DE ASOCIACIÓN

Al igual que en el Capítulo 3 será importante ver la estrategia seguida para decidir si hay o no independencia en los datos. A continuación describimos las estrategias encontradas.

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.

Esta estrategia ya apareció en el estudio inicial descrito en el Capítulo 3 aunque con poca frecuencia (8,9% para el ítem de independencia, 7,7% para el ítem de asociación inversa, 6,3% para el ítem de asociación directa y 5,8% para el ítem de asociación directa con tabla rxc). También fue descrita en el estudio de Estepa (1993) 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 (1993) y tampoco la muestra ningún sujeto en el estudio descrito en el Capítulo 3. 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 y no aparece en el Capítulo 3. 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

Chi²=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.

En el Capítulo 3, una de las estrategias correctas utilizadas por los estudiantes (S5.2) fue de este tipo, (aparece en 8 estudiantes, 1,9%) en el ítem 1 (tabla 2x2 con independencia), 5 (1,2%) en el ítem 2 (tabla 2x2 con asociación inversa), 1 (0,2%) en el ítem 3 (tabla 2x2 con asociación directa) y 3 (0,7%) en el ítem 4 (tabla rxc con asociación directa). Un caso de este tipo es:

$$\text{“} \frac{50}{96} \neq \frac{15}{39} \text{” (Estudiante 77).}$$

Estrategias incorrectas:

EI.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 (1995). Esta es una estrategia de nivel 2 en la clasificación de Pérez Echeverría (1990) y también aparece como estrategia SI2.2. en el estudio inicial descrito en el Capítulo 3 en algunos estudiantes (5.8% en el ítem 1, 14,7% en el 2, 19,1% en el 3 y 1,9% en el ítem 4). Un ejemplo es como sigue:

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

EI.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). Esta estrategia (SI5.7) se encontró en el estudio descrito en el Capítulo 3 en todos sus ítems, con una frecuencia mínima

de 12 (2,9%) estudiantes y máxima de 24 (5,8%). 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).

EI.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 (1993) y supone falta de comprensión de la necesidad de trabajar con frecuencias relativas al comparar distribuciones. Se encontró en el estudio de teorías intuitivas del Capítulo 3 (SI4.5) en todos sus ítems, con una frecuencias que variaba entre 3 (0,7%) y 6 estudiantes (1,4%), según ítem. 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. Esta estrategia no fue encontrada en el Capítulo 3. Por ejemplo:

$$50/65=0,76$$

$$96/65=0,7$$

$$15/65=0,23$$

$$39/65=0,6$$

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

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

Al observar las estrategias empleadas para llegar al juicio de asociación, se puede ver una mayoría de estrategias correctas, utilizadas por un total de 72 estudiantes (76,5%), resultados mucho mejores que los obtenidos antes de la enseñanza, donde ninguno de los ítems alcanzó el 15% de estrategias correctas (Capítulo 3), de manera que se han superado la gran mayoría de los errores cometidos anteriormente. Únicamente han empleado estrategias erróneas 5 estudiantes después de la enseñanza (5,4%), aunque hay que tener en cuenta que 14 (14,8%) no han respondido. Además aparecen estrategias formales, y no únicamente informales como en el Capítulo 3.

Vemos un resumen a continuación (Tabla 6.9.3.1). Se han abandonado prácticamente las estrategias de nivel 4 o inferior en la clasificación de Pérez Echeverría (1990) que eran mayoría en el estudio inicial de estrategias.

Tabla 6.9.3.1. Frecuencias y porcentajes de respuestas en el problema D, apartado c

	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

Síntesis de resultados en el problema D

En este problema todos los estudiantes han respondido al apartado (a) donde se les pedía una probabilidad condicional, aunque sólo el 55% dan la respuesta correcta. Por otro lado, aunque en los apartados b y c, donde se les pedía un juicio de asociación y su justificación 14 (14,8%) no han respondido, la mayoría (72%) detecta la pequeña asociación y este juicio, a diferencia de los realizados en el estudio inicial, es basado en una estrategia correcta (76,7% de los casos).

Generalmente esta estrategia es formal, basada en la comparación de las frecuencias observadas y esperadas en caso de independencia; en otros casos se usan estrategias intuitivas, pero de nivel 5 en la clasificación de Pérez-Echeverría (1990) y por tanto, correctas.

Concluimos que la enseñanza recibida ha contribuido en gran medida a mejorar la competencia de los estudiantes en la realización de juicios de asociación.

6.10. RESULTADOS EN EL PROBLEMA E: JUICIO DE ASOCIACIÓN Y JUSTIFICACIÓN

Problema E. En un estudio se analiza la relación entre la dieta pobre en calcio y el desarrollo de osteoporosis. Se obtienen los siguientes resultados

Dieta	Desarrolla osteoporosis	
	Si	No
Pobre en calcio	58	62
Rica en calcio	22	258

- ¿Hay asociación entre el tipo de dieta (rica o pobre en calcio) y el desarrollo de osteoporosis?
- ¿Es la relación directa o inversa?
- ¿Es fuerte o moderada?
- Utiliza algún coeficiente de asociación que recuerdes para justificar tus respuestas
- Si no recuerdas los coeficientes, da la justificación usando cualquier propiedad que hayas aprendido en los temas de tablas de contingencia

En este problema se pide al estudiante que emita un juicio de asociación completo y lo justifique, guiándole con los sucesivos etapas que debe realizar y que explicitamos en cuatro apartados, que estudiaremos de forma independiente: El estudiante (en el apartado *a*) puede decidir realizar una comparación entre las frecuencias esperadas en caso de independencia con las frecuencias observadas, comparar los valores de las celdas de la diagonal que informa de dependencia directa, con los de la diagonal que informa de una dependencia inversa, concluyendo que si hay asociación. Esto además le permite saber el tipo de dependencia (apartado *b*), concluyendo que la relación es directa.

En los apartados *c* y *d* puede calcular inicialmente un coeficiente de asociación, como son: Phi de Pearson = 0,464, riesgo relativo por columnas = 3,7419, riesgo relativo por filas = 6,1515, razón de productos cruzados = 10,97. Como el RR y la RC, es mayor que 1, la dependencia es directa; y como, el valor Phi de Pearson está por debajo del 0,5 se considera que es moderada. También podría comparar las frecuencias observadas con las esperadas (apartado *e*), concluyendo que son muy diferentes, observándose que hay relación entre las variables. Para resolver el ítem el estudiante también podría usar uno de los procedimientos informales que se presentaron en el estudio inicial de evaluación descrito en el Capítulo 3.

6.10.1. JUICIO DE ASOCIACIÓN.

Sólo hemos considerado las correctas (estudiantes que indican que hay asociación en los datos) e incorrectas (si indican que los datos presentan independencia). En este problema, se presenta una asociación moderada y se puede

apreciar como los estudiantes en su mayoría lo han reconocido (74 estudiantes, 79,6%). Los resultados lo podemos ver en la Tabla 6.10.1.1 son bastante buenos, lo que indica buenos resultados de aprendizaje.

Al comparar con los ítems de tablas 2x2 del cuestionario de conocimientos intuitivos (Capítulo 3), vemos como en el caso de dependencia inversa el 96,1% consideraban relación, el 3,4% consideraban independencia y 0,5% no contestan, y en el caso de dependencia directa el 93,2% consideran relación, el 5,8% consideran independencia y el 1% no contestan). Los resultados parecen empeorar, a pesar de que las teorías previas no estaban en contra, pero tenemos bastantes casos de no respuesta que bajan el porcentaje de respuestas correctas. En todo caso, el porcentaje de respuestas correctas es muy alto.

Tabla 6.10.1.1. Frecuencias y porcentajes de respuestas en el ítem problema E, apartado a

	Frecuencia	Porcentaje
RC.1 Asociación	74	79,6
RI.1 Independencia	3	3,2
No responde	16	17,2
Total	93	100

6.10.2. SIGNO DE LA ASOCIACIÓN

También en este apartado hemos considerado sólo dos tipos de respuesta: correcta, si indican que la dependencia es directa e incorrecta en caso contrario. Al estudiar los resultados (Tabla 6.2.10.1) se puede apreciar como los estudiantes en su mayoría han reconocido la dependencia directa (60 estudiantes, es decir 64,5%), mostrando de nuevo un buen aprendizaje. Recordemos que en el Capítulo 3 no se les pedía a los estudiantes que distinguieran entre una asociación directa ó inversa.

Tabla 6.10.2.1. Frecuencias y porcentajes de respuestas en el problema E, apartado b

	Frecuencia	Porcentaje
RC.1 Directa	60	64,5
RI.1 Inversa	15	16,1
No responde	18	19,4
Total	93	100

6.10.3. INTENSIDAD DE LA ASOCIACIÓN

En la pregunta sobre la intensidad de la asociación (fuerte ó moderada), también hemos considerado únicamente las respuestas correctas e incorrectas. Se puede apreciar como los estudiantes están más repartidos entre las dos opciones, a pesar de que en la diagonal principal aparecen tres veces más sujetos que en la otra diagonal, por lo que este punto no ha sido tan sencillo para los estudiantes. Hay también un alto número de no respuestas. Los resultados los podemos ver en la Tabla 6.10.3.1. En el estudio inicial de estrategias no preguntamos específicamente por la intensidad, pero si pedimos estimarla.

Tabla 6.10.3.1. Frecuencias y porcentajes de respuestas en el problema E, apartado c

	Frecuencia	Porcentaje
R.1 Fuerte	24	25,8
R.2 Moderada	38	40,9
No responde	31	33,3
Total	93	100

6.10.4. JUSTIFICACIÓN DE LA ASOCIACIÓN

Estos dos apartados los analizamos conjuntamente, ya que se pidió a los estudiantes una sola respuesta. Se obtuvieron las siguientes:

RC.1. En este caso el estudiante hace los pasos pertinentes en un ejercicio típico de un contraste de independencia. Utiliza planteamientos matemáticos, comenzando con el planteamiento de las hipótesis nula y alternativa adecuadas, calcula el estadístico de contraste, y la probabilidad asociada al valor obtenido, y por último realiza una interpretación de los resultados. Es una estrategia muy elaborada, fruto del aprendizaje en la que utiliza los conceptos de frecuencia observada y esperada, contraste de hipótesis, estadístico Chi- cuadrado. Un ejemplo es el siguiente:

$$f_e$$

24	96
56	224

$$e_1 = \frac{120 \times 80}{400} = 24$$

$$e_2 = \frac{80 \times 280}{400} = 56$$

Chi-cuadrado

48,16	12,04
20,64	5,16

$$\chi^2 = \sum \sum \frac{(f_i \cdot x_{e_i})^2}{e_i} = 86 \quad \text{'' (Estudiante 19).}$$

RC.2. El estudiante calcula bien el coeficiente Phi de Pearson comprendiendo cuál es el signo que debe considerar, ya que es el resultado de una raíz cuadrada donde el estudiante tiene que tomar la decisión del signo (Ato y López, 1996). Por tanto el estudiante tiene que combinar un procedimiento formal (la utilización de un coeficiente), con un procedimiento informal (como es la comparación de las celdas que informan de asociación directa e inversa). También esta estrategia es producto del aprendizaje. Un ejemplo aparece a continuación:

“La relación es directa, porque las mayores frecuencias se encuentran en las celdas f_{11} y f_{22}

Chi-cuadrado

48,16	12,04
20,64	5,16

$$\sum = 86$$

$$\phi = \sqrt{\chi^2 / n} = \sqrt{86,01 / 400} = 0,46 \text{ ” (Estudiante 19).}$$

RC.3. El estudiante calcula correctamente al menos uno de los valores del riesgo relativo, y además el estudiante usa una notación adecuada. Este coeficiente en su interpretación nos informa de cuanto es más probable la presencia de una variable respecto de la otra, lo cual nos dice que los estudiantes están utilizando nociones de probabilidad de forma correcta y muestran un aprendizaje formal de la materia. A continuación incluimos una respuesta de este tipo

“Riesgo Relativo

$$RR_{\text{columnas}} = \frac{A/B}{A/\bar{B}} = \frac{f_{11} / f_{.1}}{f_{12} / f_{.2}} = \frac{58 \times 320}{62 \times 80} = 3,74 \text{ ” (Estudiante 19).}$$

RC.4. El estudiante calcula de forma correcta la razón de productos cruzados, dando una interpretación adecuada de los resultados del problema, por lo que utiliza nociones de razones que son necesarias en la interpretación de este coeficiente. Un ejemplo es el siguiente:

“El coeficiente de productos cruzados (RC)

$$RC = \frac{58 \times 258}{22 \times 62} = \frac{14964}{1364} = 10,97 \text{ ” (Estudiante 69).}$$

RC.5. Como el caso anterior, el estudiante calcula dos coeficientes de asociación para confirmar los resultados: el valor Phi de Pearson (RC.2) y la razón de productos cruzados (RC.4), como vemos en el siguiente ejemplo

“Razón de productos cruzados

$$RC = \frac{58 \times 258}{22 \times 62} = 10,9$$

Sabiendo que $RC > 1$ existe una asociación positiva alta

$$\phi = \sqrt{\chi^2 / n} \rightarrow \text{Coeficiente Phi de Pearson}$$

Frecuencias esperadas

$$\frac{120 \times 80}{400} = 24 \quad \frac{120 \times 320}{400} = 96$$

$$\frac{280 \times 80}{400} = 56 \quad \frac{320 \times 280}{400} = 224$$

$$\chi^2_{\text{exp}} = 48,16 + 12,04 + 20,6 + 5,16 = 85,96$$

$$\phi = \sqrt{85,96 / 400} = 0,463$$

Se trata de una asociación directa, positiva y moderada” (Estudiante 19).

RC.6. El estudiante en este caso utiliza el valor Phi de Pearson (RC.2) y el riesgo relativo (RC.3) para corroborar los resultados que obtiene. Un ejemplo se presenta en lo que sigue:

“

Dieta	Osteoporosis		Total
	Si	No	
Ri. Calcio	24	96	120
Pob. Calcio	56	224	280
Total	80	320	400

χ^2_{exp}

$$\frac{(58 - 24)^2}{24} = 48,16 \quad \frac{(62 - 96)^2}{96} = 12,04$$

$$\frac{(22 - 56)^2}{56} = 20,64 \quad \frac{(258 - 224)^2}{224} = 5,16$$

$\Rightarrow +86$

$(2-1) \times (2-1) = 1 \times 1 = 1$ Grados libertad

La relación es directa porque phi de Pearson = 0,46

La relación es fuerte $RR > 1$

$$RR = \frac{58 \times 320}{62 \times 80} = 3,74 \text{ ” (Estudiante 65).}$$

RC.7. El estudiante utiliza el valor del riesgo relativo (RC.3) y la razón de productos cruzados (RC.4) para justificar su respuesta, como vemos en el siguiente ejemplo:

$$“RR_{\text{filas}} = 6,15$$

$$RC = 10,9$$

Valores altos, directa” (Estudiante 104).

RC.8. Comparar las distribuciones condicionales por fila ó por columna. Es decir, el estudiante aplica una de las propiedades de la independencia, la cual dice que si las distribuciones condicionales, por filas o por columnas, coinciden, entonces las variables son independientes. En consecuencia están usando todas las celdas de la tabla y razonando en el nivel 5 de Pérez Echeverría (1990), además en este caso utilizan conceptos de frecuencias condicionales e independencia.

Sería una estrategia informal que también apareció en el estudio inicial (entre el 5,8% y el 8,9%), descrito en el Capítulo 3, comparar todas las distribuciones de frecuencias relativas condicionales de una variable para los distintos valores de la otra variable (SC5.1). Por ejemplo:

“Utilizando las frecuencias condicionales obtenemos que las condicionales para x son diferentes a las frecuencias para y . Haciéndolo por filas” (Estudiante 9).

RC.9. El estudiante en este caso, calcula las frecuencias esperadas y las compara con las frecuencias observadas. En este caso, el estudiante aplica una de las propiedades de la independencia, donde se dice que si las frecuencias esperadas en caso de independencia, coinciden con las frecuencias absolutas, entonces las variables son independientes. Utilizan conceptos de frecuencias esperadas e independencia; es una estrategia formal fruto del aprendizaje.

“He calculado las frecuencias esperadas” (Estudiante 85).

Respuestas parcialmente correctas:

RP.1. El estudiante puede no hacer ningún tipo de cálculo, sino comparar los valores de las celdas que informan de una dependencia directa (a y d), con los valores de las celdas que informan de una dependencia inversa (b y c), llegando a la misma conclusión que en casos anteriores. Esta estrategia aparece de modo informal en el Capítulo 3, con el nombre *“Comparar la suma de los valores en las diagonales”* (SP4.1), su frecuencia en el capítulo 3 es entre 0% y el 1,2%. Fue descrita por Shaklee y Tucker (1980), Allan y Jenkins (1983) y aparece en Estepa (1993), Un ejemplo es:

$$(58+258)-(22+62)=232$$

Hay asociación positiva moderada y positiva analizando las diagonales. La relación en la diagonal principal es mayor que en la diagonal inversa” (Estudiante 29).

Como se observa en la Tabla 6.10.4.1 38,8% de los estudiantes fue capaz de calcular un coeficiente de asociación correcto; en ocasiones más de uno. El 17,2% calcula el estadístico Chi-cuadrado y 15,1% el Phi de Pearson, a pesar de no disponer ese día de programa de cálculo y ser más complejos que el riesgo relativo, o la razón de productos cruzados, que fueron calculados por pocos estudiantes. Interpretamos que su significado fue menos claro para los estudiantes que los dos primeros, posiblemente por no tener un valor máximo posible.

Tabla 6.10.4.1. Frecuencias y porcentajes de respuestas en el problema E, apartado d y e

Apartado			Frecuencia	Porcentaje
d	RC.1	Chi-cuadrado	16	17,2
	RC.2	Phi de Pearson	14	15,1
	RC.3	Riesgo relativo	1	1,1
	RC.4	Razón de productos cruzados	2	2,1
	RC.5	Phi de Pearson y razón de productos cruzados	1	1,1
	RC.6	Phi de Pearson y riesgo relativo	1	1,1
	RC.7	Riesgo relativo y razón de productos cruzados	1	1,1
e	RC.8	Comparar las distribuciones condicionales	12	12,9
	RC.9	Comparar frecuencias absolutas con frecuencias esperadas	2	2,1
	RP.1	Comparar la suma de las diagonales	2	2,1
No responde			41	44,1
Total			93	100

Fueron también bastantes (12,9%) quienes calcularon y compararon las distribuciones condicionales, es decir, trabajaron con estrategias intuitivas, similares a las encontradas en el estudio del Capítulo 3, aunque a nivel 5 de Pérez Echeverría (1990).

En este apartado 41 estudiantes no respondieron (44,1%) un porcentaje de no respuesta mucho mayor que en los apartados anteriores, donde los estudiantes ya habían en su mayoría dado respuestas correctas. Por ello pensamos que la falta de respuesta en este apartado fue debida más a una falta de capacidad de argumentación que a no haber utilizado un procedimiento adecuado, pues en dicho caso no se hubiesen obtenido resultados correctos con tan alto porcentaje en los apartados anteriores.

También puede ser que el alto porcentaje de no respuestas en los apartados c, d ó e puede ser debido al contrato didáctico, el estudiante puede razonar que si en un problema de 4 apartados, ya ha hecho dos y, probablemente correctos, ya ha obtenido puntuación en este problema, en consecuencia, si no está seguro, no responde. Por otro lado, incluso teniendo en cuenta la no respuesta, el porcentaje de estrategias correctas (en su mayoría formales, fruto del aprendizaje) es mucho mayor que en el estudio inicial de evaluación.

Síntesis de resultados en el problema E

En este problema el 79,6% de los estudiantes da un juicio de asociación correcto en el apartado (a); únicamente 3 casos (3,2%) consideran erróneamente que existe independencia y el resto no responde. Por otro lado, en el apartado b, donde se pedía el signo de la asociación, un 64,5% de estudiantes identifica correctamente la asociación

directa, un 16,1% incorrectamente la asociación inversa y el resto no responde. Respecto a la intensidad de la asociación (apartado c), los estudiantes están más repartidos entre las dos opciones: asociación fuerte (25,8%) y moderada (40,9%), a pesar de que la diagonal principal aparecen tres veces más sujetos que en la diagonal secundaria.

En la justificación de la asociación (apartados (d) y (e)), se obtienen un 53,8% de respuestas correctas, siendo la estrategia más utilizada, la realización del contraste Chi-cuadrado (17,2%) y el cálculo del coeficiente Phi de Pearson (15,1%). Este apartado es el que obtiene el mayor número de no respuestas (44,1%), pero casi todos los que responden lo hacen de forma correcta y con un nivel 5 en la clasificación de Pérez-Echeverría (1990).

Concluimos una mejora por parte de los estudiantes en todas las fases del proceso de resolución de este tipo de problemas, una muestra de que el proceso de enseñanza realizada resulta positivo.

6.11. SINTESIS DE RESULTADOS EN LOS PROBLEMAS ABIERTOS

De los 104 estudiantes que participaron en la enseñanza, un total de 82 realizaron los cinco problemas abiertos analizados en este capítulo. Puntuando un punto por cada apartado correcto (cada problema tiene un número diferente de apartados) podemos obtener una puntuación total de hasta 16 puntos: 4 puntos del problema uno y tres en cada uno de los restantes. En la Tabla 6.11.1 presentamos los estadísticos descriptivos de estas calificaciones, obteniendo unos resultados satisfactorios, destacando que la media está por encima del valor teórico (8), y que el valor máximo obtenido por los estudiantes está muy cercano al valor máximo, aunque ninguno llegó a completar todos los apartados de los cinco problemas correctamente.

Hay, en consecuencia, evidencias del aprendizaje también en los problemas abiertos en que los estudiantes han de aplicar los conocimientos adquiridos teóricamente, que también fueron buenos, como se vio en el Capítulo 5.

Tabla 6.11.1. Estadísticos descriptivos

	Mínimo	Máximo	Media	Desv. típica
Nota	4	15	9,46	2,842

En las Figuras 6.11.1 y 6.11.2, podemos observar gráficamente los resultados obtenidos en el histograma y el diagrama de caja, apreciándose la mejora del conocimiento de los estudiantes por el bloque tan numeroso de alumnos (52) con puntuaciones 9 o 10, es decir, sobre la media teórica y sólo 22 de ellos por debajo de la media teórica, incluso llegando a un grupo importante que completa casi todos los problemas y apartados. Además, se puede apreciar en el histograma como los datos parecen distribuirse simétricamente alrededor de las modas. La variabilidad de la puntuación (de 4 a 15 puntos) indica buena capacidad de discriminación del cuestionario.

Respecto al gráfico de caja, de nuevo se observa una distribución bastante simétrica, sin valores atípicos (ningún estudiante se comportó de forma atípica en la resolución del total de problemas. La mediana (9) es algo inferior a la media, por lo que 50% resuelven 9 o más apartados correctamente y un 25% 11 apartados o más. El 50% de los casos centrales resuelve entre 7 y 11 problemas, es decir un poco menos o más de la mitad de los propuestos, tirando un poco por encima de la media.

Figura 6.11.1. Histograma de las puntuaciones totales

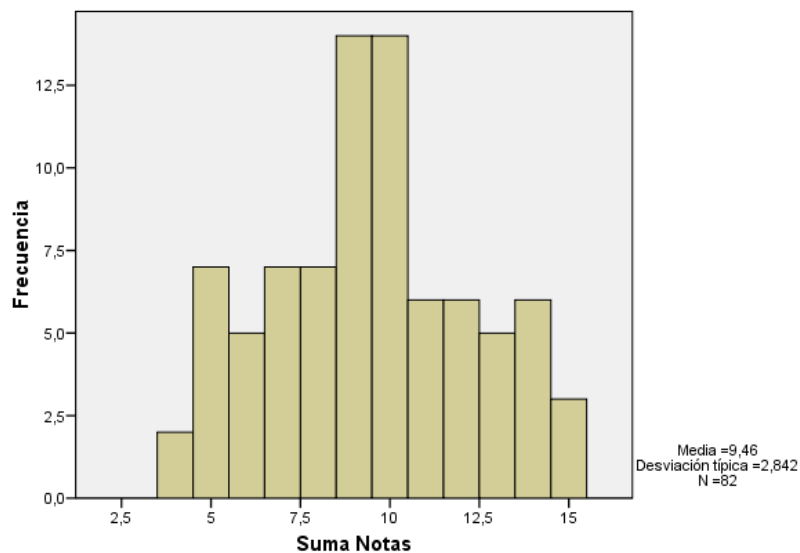


Figura 6.11.2. Diagrama de la caja de las puntuaciones totales

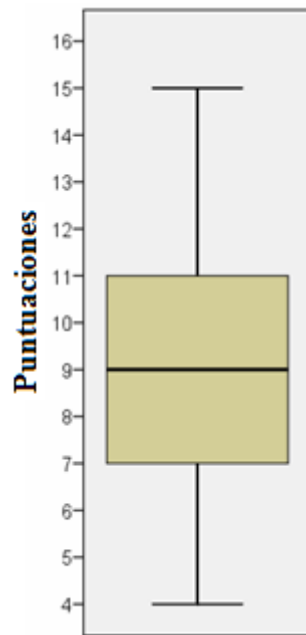


Tabla 6.11.3. Porcentajes de respuestas correctas por apartado

Problema y Apartado	Respuestas correctas	Total respuestas	Porcentaje
D3. Estrategia en juicio de asociación	76	94	0,8085
E1. Juicio de asociación sin apoyo de Excel	74	93	0,7957
B3. Calculo/interpretación riesgo relativo	68	92	0,7391
D2. Juicio de asociación sin apoyo de Excel	68	94	0,7234
E2. Signo de la asociación	60	93	0,6452
A2. Cálculo del estadístico Chi-cuadrado y p -valor con Excel	58	92	0,6304
D1. Calculo probabilidad condicional	56	94	0,5957
B3. Cálculo e interpretación razón de productos cruzados	54	92	0,587
A3. Aceptación/Rechazo hipótesis homogeneidad	54	92	0,587
E3. Intensidad de la asociación	52	93	0,5591
A1. Plantear las hipótesis del contraste Chi-cuadrado	43	92	0,4674
C2. Cálculo e interpretación V de Cramer	41	92	0,4457
A4. Interpretación de resultados de contraste homogeneidad	40	92	0,4348
C1. Cálculo e interpretación coeficiente de Contingencia	36	92	0,3913
B1. Cálculo e interpretación Phi de Pearson	33	92	0,3587
C3. Cálculo e interpretación Lambda de Goodman y Kruskal	29	92	0,3152

Por otra parte, para estudiar la dificultad relativa de los problemas ordenamos el porcentaje de respuestas correctas para cada apartado de cada problema (Tabla 6.11.3). Se puede apreciar que los tres apartados con mayor dificultad fueron la interpretación de los coeficientes de la asociación. Respecto a la interpretación del coeficiente Lambda de Goodman y Kruskal la razón de la dificultad es que su interpretación es diferente al

resto de los coeficientes. En el coeficiente Phi de Pearson la dificultad podría explicarse por los errores al considerar el signo de la raíz cuadrada y en el coeficiente de contingencia de Pearson, el hecho de que no tenga un máximo fijo, sino que dependa del número de filas y columnas.

Fue muy sencillo para los estudiantes el elegir una estrategia correcta en juicios de asociación (D3), llevar a cabo dichos juicios de asociación (E1) (D2), cálculo e interpretación de riesgo relativo (B3) y signo de la asociación (E2) así como cálculo del estadístico Chi-cuadrado y valor p con Excel (A2).

El resto de los apartados tuvo una dificultad moderada pues el porcentaje de respuestas correctas varía entre el 43% y 60%.

6.12. RELACIÓN ENTRE CONOCIMIENTO TEÓRICO Y RESOLUCIÓN DE PROBLEMAS

Para finalizar el estudio en este apartado se pone en relación las respuestas al cuestionario descrito en el Capítulo 5 y las soluciones a los problemas abiertos analizadas anteriormente, con la finalidad de analizar la influencia de los conocimientos teóricos adquiridos en el tema, sobre la competencia en la resolución de problemas.

Tabla 6.11.3. Puntuación media total en cuestionario (Capítulo 5) de los estudiantes según dan respuesta correcta o no correcta a cada apartado

Apartado	Solución en el problema		t	Sig.
	Correcta	No correcta		
D3. Estrategia en juicio de asociación	54,69	54,00	0,179	0,858
E1. Juicio de asociación sin apoyo de Excel	55,05	53,33	0,495	0,622
B2. Calculo/interpretación riesgo relativo	54,65	53,09	1,166	0,253
D2. Juicio de asociación sin apoyo de Excel	54,83	55,08	-0,123	0,903
E2. Signo de la asociación	55,68	53,00	2,042	0,049*
A2. Cálculo del estadístico Chi-cuadrado y p -valor con Excel	54,04	55,68	-1,059	0,293
D1. Calculo probabilidad condicional	55,07	53,65	1,071	0,287
B3. Cálculo e interpretación razón de productos cruzados	54,98	53,21	0,975	0,333
A3. Aceptación/Rechazo hipótesis homogeneidad	54,53	54,91	-0,235	0,815
E3. Intensidad de la asociación	54,58	53,96	0,413	0,681
A1. Plantear las hipótesis del contraste Chi-cuadrado	55,59	53,62	1,368	0,175
C2. Cálculo e interpretación V de Cramer	56,17	53,13	1,907	0,061
A4. Interpretación de resultados de contraste homogeneidad	54,92	54,12	0,464	0,644
C1. Cálculo e interpretación coeficiente de Contingencia	54,27	55,64	-0,869	0,388
B1. Cálculo e interpretación Phi de Pearson	54,70	54,53	0,113	0,910
C3. Cálculo e interpretación Lambda de Goodman y Kruskal	55,62	53,79	1,133	0,261

* Diferencias estadísticamente significativas

En la Tabla 6.11.3 se presentan los resultados del contraste t de diferencias de medias en muestras independientes en la puntuación total en el cuestionario de los alumnos que responden correctamente o no (incluyendo en estas categorías las respuestas incorrectas y parcialmente correctas) en cada apartado de los problemas. Recordemos que la puntuación total varía de 0 a 78 y se obtuvo un valor medio de 54.49.

Podemos observar como únicamente se encuentran diferencias estadísticamente significativas entre los estudiantes que responden en el problema E en el segundo apartado, referido al signo de la asociación entre las variables, lo que quiere decir que los que respondieron bien este apartado tienen mejores conocimientos de la teoría. En líneas generales, los estudiantes que respondieron a los problemas correctamente o no, obtuvieron puntuaciones similares, en torno a la media en el cuestionario teórico. Concluimos que la adquisición de conocimientos teóricos fue relativamente sencilla, pero no siempre se transfiere a la resolución de problemas.

6.13. CONCLUSIONES

En este capítulo se han analizado las respuestas a los problemas abiertos propuestos a los estudiantes, tres de los cuáles (problemas A, B y C) fueron realizados durante la sesión de practicas, con ayuda del ordenador, y otros dos problemas (problemas D y E) resueltos durante las sesiones de evaluación (junto a los cuestionarios tipo test vistos en el Capítulo 5), donde los estudiantes trabajaron sin ningún tipo de ayuda. El análisis a priori de los problemas indicó que se han tenido en cuenta la mayor parte de los objetos incluidos en la enseñanza, asegurando la validez de contenido al proceso de evaluación.

Nuestros objetivos al proponer a los estudiantes estos problemas fueron profundizar en la evaluación de su comprensión y aplicación de los objetos matemáticos incluidos en la enseñanza, observar las estrategias de resolución de los problemas, estudiar si adquirieron dominio del programa realizado en Excel sobre tablas de contingencia y analizar sus razonamientos. A continuación se hace un resumen de las principales conclusiones obtenidas.

Competencia lograda en la resolución de los problemas

Un primer resultado cuantitativo se refiere al número medio de apartados correctos en el conjunto de todos los problemas, que fue mayor que el valor medio

teórico esperado, evidenciando la adquisición de las competencias pretendidas. En promedio, cada estudiante resolvió un número de apartados mayor (9,46) que el valor medio teórico (8), y la mayoría respondieron correctamente al menos a la mitad de los subítems. La variabilidad es pequeña, al comparar el valor de la desviación típica respecto a la media, y algunos estudiantes alcanzan 15 puntos de 16. Por otra parte, aunque el porcentaje de estudiantes que resuelven cada ejercicio de forma completa correctamente es mejor, los porcentajes de los que lo logran son bastante satisfactorios.

Además, hay que mencionar que la dificultad de los problemas fue variada, encontrándose una variabilidad desde el 31,5% de respuestas correctas, hasta un 80,8%. El análisis cuantitativo de los porcentajes de respuestas en cada apartado de cada problema, permitió identificar los puntos de mayor dificultad, relacionados con la interpretación de algunas de las medidas de asociación, en particular si no tienen un máximo fijo o si el estudiante ha de deducir su signo.

Juicios de asociación: estrategias y estimación de su intensidad

Además de estos resultados cuantitativos, se ha hecho un análisis cualitativo detallado de las respuestas en cada apartado de todos los problemas abiertos, comparando con la investigación previa. También se proporcionan muchos resultados originales, como son la clasificación de tipos de respuestas en el estudio de los coeficientes de asociación, donde no se encontraron investigaciones previas. Entre otros resultados, resaltamos los siguientes:

- En los juicios de asociación (problema E) la mayoría de estudiantes es capaz de utilizar una estrategia correcta, la mayoría de las cuáles son formales, fruto del aprendizaje. Aunque en algunos casos, se trata de estrategias intuitivas vistas anteriormente, (Capítulo 3) generalmente son estrategias de nivel 5 en la clasificación de Pérez- Echeverría (1990). Puesto que en el estudio descrito en el Capítulo 3 se pudo observar una variedad de estrategias incorrectas y tan sólo entre el 11,9% y 18,9% de nivel 5, se aprecia claramente el efecto de la enseñanza sobre la competencia en los juicios de asociación.
- No se ha encontrado un efecto de las teorías previas sobre la asociación en los problemas planteados posteriores a la enseñanza, mientras esta influencia fue importante en los problemas propuestos en el Capítulo 3. Deducimos, por tanto una influencia positiva de la enseñanza para lograr que los alumnos usen con preferencia los datos, frente a sus teorías previas en los juicios de asociación.

- Hay también una gran mejora en la precisión de la estimación de la asociación, en cuanto los estudiantes calculan mediante algún procedimiento formal algún coeficiente, que, en un porcentaje apreciable logran interpretar correctamente asociación. Por el contrario, en el Capítulo 3, se observaba una sobrestimación, que era especialmente importante cuando las teorías previas estaban a favor. La enseñanza ha contribuido también a superar este sesgo.
- Mientras en el estudio descrito en el Capítulo 3 más de un 70% de los estudiantes operaban con frecuencias absolutas, con la enseñanza parece haberse corregido esta tendencia, pues se observa que saben distinguir cuándo utilizar cada tipo de frecuencias y utilizan generalmente frecuencias relativas o bien frecuencias esperadas en caso de independencia.
- Los estudiantes logran competencia en el manejo del software proporcionado, y en resolver problemas con este recurso, que se deduce del cálculo correcto que realizan de las medidas de asociación en los problemas B y C. Los únicos errores encontrados fueron errores al construir la tabla de contingencia, en el cálculo de grados de libertad y no considerar el número de filas y columnas obtener el valor máximo de algunos coeficientes en el problema C.
- También hubo una gran mejora en el cálculo de probabilidades condicionales a partir de tablas de contingencia en el problema D, ya que nuestros resultados mejoraron los obtenidos por todas las investigaciones previas con las que se compara (Díaz, 2007; Estrada y Díaz, 2007; Contreras, 2011). Nuestras respuestas (59,6%) superan en porcentaje a todas estas investigaciones.
- Los estudiantes han aprendido la fórmula de cálculo de los grados de libertad en su mayoría y apenas aparece el error citado por Alvarado (2007) y Olivo (2008) consistente en no restar el número de restricciones en el cálculo (ejercicio A, apartado b).

Como puntos negativos, destacamos que, aunque los estudiantes muestran un buen aprendizaje de los objetos matemáticos relacionados con la tabla de contingencia, han tenido dificultad al interpretar las soluciones de los problemas en el contexto de dicho problema cuando se les pide. Como ya se indicó en este apartado, este es el último paso en el proceso de modelización según Henry (1997); el contrato didáctico con el estudiante implica que en pocas ocasiones los profesores llegan a este último

paso, a pesar de su importancia. Este sería un punto para profundizar en la futura enseñanza del tema.

Conflictos semióticos identificados

El análisis de las respuestas abiertas a los problemas permitió también identificar algunos conflictos semióticos, que clasificamos en lo que sigue en varios apartados.

Conflictos semióticos relacionados con conceptos o propiedades

- Los estudiantes intercambiaron las hipótesis alternativas y nula, un error descrito por Vallecillos (1994) y Chow (1996) (ver IR1 e IR3, en el apartado a, del problema A). Estos estudiantes confunden las propiedades de estas dos hipótesis: La hipótesis nula se establece para ser rechazada y determina la distribución muestral del estadístico (Batanero, 2000); no se puede trabajar directamente con la hipótesis alternativa, pues no es una hipótesis puntual.
- En algunas respuestas los estudiantes confunden el estadístico y el parámetro (IR2, en el apartado a, del problema A), error descrito entre otros por Schuyten (1991). Mientras que la distribución de probabilidad de una variable depende de algunos parámetros (normalmente desconocidos y constantes), los estadísticos se calcula a partir de los datos de la muestra y son variables aleatorias, aunque el valor para el muestra particular es fijo y conocido (Batanero, 2000). Por ello no tiene mucho sentido para establecer las hipótesis en términos de los estadísticos (ya que su valor es conocido en la muestra particular). Además, en esta respuesta los estudiantes no recuerdan que el valor p es una probabilidad y por lo tanto su valor no puede ser superior a 1.
- También se encontró una cierta confusión entre el p-valor y el nivel de significación, dos conceptos que, de acuerdo con Morrison y Henkel (1970), Menon (1993), Vallecillos (1994) y Williams (1997) son particularmente mal entendidos. El nivel de significación se define como la probabilidad de falso rechazo de una hipótesis nula y es un valor constante que se establece a priori (antes de recopilar los datos) por el investigador. El valor p se define como la probabilidad de observar el valor empírico del estadístico o un valor más extremo, dado que la hipótesis nula es verdadera y varía de una muestra a otra. La confusión entre estos dos conceptos que fue encontrado en el IR4 respuesta en el apartado b) del problema A.

- Otros estudiantes confundieron las regiones de aceptación y rechazo como era visible en las respuestas incorrectas en el apartado c), así como en las respuestas IR1 e IR2 en el apartado d) del problema A. Estos errores también descritos por Haller y Kraus (2002) implican una mala interpretación del razonamiento lógico detrás de las pruebas estadísticas. Según Batanero (2000), este malentendido se debe al hecho de que, al parecer, la estructura formal de las pruebas estadísticas es similar a la de la prueba por reducción al absurdo, sin embargo, existen diferencias fundamentales entre estos dos tipos de razonamiento que no siempre son así comprendidas por los alumnos.
- Las dos hipótesis no cubre siempre el espacio paramétrico (véanse las respuestas PR3, IR3 y IR6, en el apartado a, del problema A); aquí de nuevo el estudiante falla en la comprensión del razonamiento lógico en los test estadísticos, donde la hipótesis nula es el complemento lógico de la alternativa. Este error también fue descrito en Vallecillos (1994).
- Confundir una variable y sus valores (IR4 en el apartado a, del problema A) o confusión por el número de variables y el número de valores (IR5 en el apartado a, del problema A).
- Confundir las frecuencias relativas dobles con las frecuencias esperadas, al intentar aplicar una de las propiedades de independencia (ver EI.2, apartado c, problema D). Este conflicto aparece también en Díaz (2007) y Contreras (2011).
- El estudiante compara frecuencias marginales creyendo en forma errónea, que si las frecuencias marginales coinciden, entonces hay independencia (ver EI.3, apartado c, problema D). Esta estrategia también fue encontrada en Estepa (1993) y supone falta de comprensión de la necesidad de trabajar con frecuencias relativas al comparar distribuciones.
- En algunos casos se realizan interpretaciones erróneas de los coeficientes de asociación, confundiendo la dependencia directa con la dependencia inversa (RP.3 y RP.4, apartado a, problema B). O bien se deja la interpretación en blanco, sin dejar claro (RP.5 y RP.6, apartado a, problema B; RP.1, apartado b, problema B; RP.1, apartado c, problema B; RP. 2, apartado b, problema C; RI.1, apartado b, problema C; RP.1, apartado c, problema C). En otros casos consideran independencia, en lugar de una asociación moderada ó moderada-baja (RP.1, apartado, problema C).
- Algunos estudiantes, muestran un conflicto específico en la interpretación de la

intensidad (ver RP.1 y RP.2, apartado a, problema C). También en la interpretación de Lambda de Goodman y Kruskal, considerando una asociación moderada, en lugar de en términos de predicción de error (ver RP.2, apartado c, problema C).

- Otro conflicto, aparece relacionada con el concepto de “razón”, ya que hay estudiantes que no interpretan de forma eficiente el coeficiente relacionado con este concepto (RP.1 y RP.2, apartado c, problema B).

Conflictos semióticos involucrados en los procedimientos

- Al afirmar sus hipótesis, parte de los alumnos confunden la prueba de homogeneidad y la prueba de independencia (las respuestas IR1 y IR4, en el apartado a, del problema A). Aunque en la enseñanza estudiaron la distinción entre ambas pruebas (como la mayoría de los libros de texto las estadísticas lo hacen), estos estudiantes relacionaron “independencia” y “Chi-cuadrado”. Según lo sugerido se necesita un tiempo para que los estudiantes acepten un uso más general de esta distribución, por lo que será importante presentarles más ejemplos de estas situaciones, que son frecuentes en la investigación, como por ejemplo, al estudiar el acuerdo de los evaluadores.
- Cálculo incorrecto de los grados de libertad en la distribución Chi-cuadrado (respuestas IR1, IR3, en el apartado b, del problema A). Este error también se observó en investigaciones previas por Alvarado (2007) y Olivo (2008), que sugieren la necesidad de encontrar la manera de aclarar a los estudiantes este concepto, que es de gran importancia para la teoría estadística moderna y se utiliza en varias distribuciones muestrales importantes.
- El estudiante calcula los coeficientes correspondientes a la asociación de tablas 2x2 (coeficiente Phi de Pearson, riesgo relativo y razón de productos cruzados), en lugar de los adecuados para tablas rxc (coeficiente de contingencia de Pearson, V de Cramer y Lambda de Goodman y Kruskal). Estos estudiantes tienen un conflicto al no saber cuándo aplicar los coeficientes de asociación para tablas 2x2, y para tablas rxc (ejemplos encontrados en el problema C: RI.2, apartado a; RI.3, apartado b; RI.2, apartado c).
- Algunos estudiantes no calculan el valor Lambda de Goodman y Kruskal, debido a que consideraron que no es una medida de asociación, pues en los apuntes se presentaba como una medida basada en la reducción proporcional del error (RI.3,

apartado c, problema C).

- No usas todas las celdas de la tabla (ver EI.1, apartado c, problema D). Generalmente se usan dos celdas (presencia-presencia y ausencia-ausencia) que informarían que la asociación en la tabla es directa, según Inhelder y Piaget (1995). Esta es una estrategia de nivel 2 en la clasificación de Pérez Echeverría (1990).
- Muchos estudiantes en la aplicación de los coeficientes de asociación confunden los signos de las raíces cuadradas, que han de dar como solución en las formulas de los mismos (las respuestas RP.1, RP.2, RP.4 y RP.5, apartado a, problema B). El conflicto puede darse en la utilización de raíces cuadradas (Roach, Gibson y Weber, 2004); otro posible explicación es una mala utilización de los procedimientos informales como la comparación de las diagonales (Inhelder y Piaget, 1955) para deducir el signo.
- Algunos estudiantes muestran conflictos en el cálculo de formulas de los coeficientes de asociación del problema B (ver RP.2, apartado a; RP.3, apartado b; RP.3, apartado c), y del problema C (RI.1, apartado a). Algunas formulas son el calculo del valor máximo posible (RP.3, RP.4 y RP.5, apartado a, problema C); en este caso puede surgir el conflicto por no considerar el número de filas y columnas (RP.6, apartado a, problema C; RP.3, apartado b, problema C).
- Otro conflicto de cálculo aparece al intentar calcular las frecuencias condicionales, para aplicar una de las propiedades de la independencia, donde el alumno divide las cuatro frecuencias absolutas por una frecuencia absoluta marginal (RI.4, apartado c, problema D).

Conflictos Semióticos involucrados con el cálculo de probabilidades

- En algunos casos, los alumnos confunden el tipo de probabilidad que se necesita, calculando la probabilidad conjunta, en lugar de la probabilidad condicional (RI.1, apartado a, problema D). Pollatsek, Well, Konold y Hardiman (1987) consideran que este error surge por la falta de precisión del lenguaje verbal de los enunciados de los problemas cuando expresamos probabilidades condicionales. Asimismo Einhorn y Hogarth (1986), observaron que algunos estudiantes, confundían la interpretación de estas dos probabilidades, y la mitad de los sujetos del estudio de Ojeda (1995) interpretaron la intersección como condicionamiento. También aparece en el estudio de Contreras (2011).

- Otra confusión es calcular la probabilidad simple, en lugar de la probabilidad condicional (RI.2, apartado a, problema D). Este conflicto aparece también en otras investigaciones (Estrada y Díaz, 2007; Contreras, 2011).
- Algunos estudiantes calculan una probabilidad condicional diferente de la que se le pide (RI.3, apartado a, problema D). Batanero, Estepa, Godino y Green (1996), observaron la confusión entre las dos frecuencias condicionales relacionadas con una misma celda de la tabla de contingencia, lo que en nuestro caso podría producir este error. Como en otras respuestas anteriores Falk (1986) considera como explicación de este error, que denomina falacia de la condicional transpuesta la imprecisión del lenguaje cotidiano. También aparece en Contreras (2011).
- En el cálculo de probabilidades condicionales, algunos estudiantes dividen dos frecuencias absolutas, que no llevan a una probabilidad sino a una razón; es decir, confunde estos dos conceptos (RI.4, apartado a, problema D). En la investigación de Contreras (2011) aparece un sujeto que realiza esto mismo.
- En el cálculo de probabilidades condicionales, algunos estudiantes confunde la fórmula de la probabilidad condicional, intercambiando el numerador con el denominador. Esto le lleva a obtener una probabilidad superior a uno, por lo que no es consciente de los axiomas de la probabilidad (RI.5, apartado a, problema D). Contreras (2011) también describe 2 casos (1,1%).

También hay que mencionar que el número de no respuestas aumento considerablemente respecto a los resultados del Capítulo 3, pero a pesar de ellos las respuesta correcta no sólo superan mucho a los del estudio realizado en el Capítulo 3 sino que están por encima del 50% en la mayoría de los apartados de los problemas.

No se observa una relación clara entre los conocimientos teóricos mostrados en el cuestionario descrito en el Capítulo 5 y la realización correcta de los problemas; posiblemente la explicación es que los resultados en el cuestionario fueron muy buenos, ya que los estudiantes sobrepasan en general, el valor medio teórico del número de respuestas correctas. Otra conclusión es que fue más sencillo adquirir el conocimiento teórico, pero este no siempre se transfiere a conocimiento práctico utilizado en la resolución de los problemas.

Por último, destacamos que todos estos resultados de la evaluación del aprendizaje, junto con los vistos para preguntas cerradas (Capítulo 5), es una aportación

de este trabajo de investigación de destacado valor para futuras planificaciones de la enseñanza, así como para la evaluación de los aprendizajes adquiridos y para el inicio de otras investigaciones.

CHAPTER 7.

CONCLUSIONS

- 7.1. Introduction
- 7.2. Conclusions regarding the general aims
- 7.3. Conclusions regarding the hypotheses
- 7.4. Study contributions and limitations
- 7.5. Ideas for future research

7.1. INTRODUCTION

Along this document we have presented a series of studies related to the understanding of contingency tables by psychology students, which are based on the Onto-semiotic approach for mathematics education (Godino, 2002b; Godino, Batanero & Font, 2007), which is the theoretical framework that has supported our research.

Firstly, we carried out an extensive study of the intuitive understanding and estimation of association in contingency tables by a sample of 414 students from three universities, before the formal study of the subject (Chapter 3). This study partly replicates another study conducted by Estepa (1993) with high school students, although the items used by the author were slightly modified and his analysis of responses was expanded. In addition, we designed and tested a study process of contingency tables, which was designed for psychology students, and which took into account the analysis of the reference meaning (Chapter 1), the main challenges identified in the previous evaluation study (Chapter 3) and in previous research (Chapter 2). This study process is described in detail in Chapter 4 and the learning outcomes of the participating students are assessed in chapters 5 and 6 with instruments that were specifically built for this research.

Since we analyzed thoroughly the conclusions reached on their specific objectives and hypotheses for each of these studies, in this final chapter, to not being repetitive, we will only present the general conclusions which follow from the set of studies carried out. These findings are discussed in relation to the overall objectives of the work, the contributions and limitations thereof and some possible research lines to expand our results.

7.2. CONCLUSIONS REGARDING THE GENERAL AIMS

In the first chapter we described the following general aims for this research:

O1: Conducting a semiotic analysis of the mathematical object "contingency tables", using the types of mathematical objects (elements of meaning) considered in the Onto-semiotic approach. The purpose was to describe the problem fields, concepts, properties, procedures, language and arguments related to contingency tables.

To meet this goal, we firstly develop in Section 1.3 the theoretical framework used in this research. Later in Section 1.4 we make a detailed study of the mathematical object "contingency tables", which was published in Cañadas, Batanero and Arteaga (2010). We begin with the identification of the problem fields from which this object arises, and then describe the different concepts, properties, procedures, language and arguments required in solving these problems.

The main *problem-situations* are classified into five types: (a) summarizing information from a set of bivariate observations, (a) computing probabilities associated with specific values of one or both variables summarized in a contingency table, (b) homogeneity test, to decide whether the distribution of a qualitative variable in different populations can be considered equivalent or not, (c) independence test, to analyze whether the association between two variables in a contingency table is statistically significant, and (d) measuring the strength of association between variables and the sign of association, when possible.

Basing on the above identification, we subsequently designed a teaching process that was experimented with psychology students. A lesson was devoted to the first two problem fields, a second lesson to the concept of independence and its properties, the third to study the independence and homogeneity tests and the last lesson to study the measures of association and their properties. Each of these problem fields is associated with some epistemic configurations (Godino, 2002b; Godino, Batanero & Roa, 2005; Font, Godino & D'Amore, 2007) that vary in richness and complexity, as noted in the following mathematical objects implicit in the work with these problems:

- *Language:* In addition to a variety of verbal expressions, both from common language and specific for mathematics, we identified the following types of graph: bar graph, attached and stacked bar graph, mosaic graph, three-dimensional histogram. Symbolic language with a wide variety of formulas and algebraic

expressions, and tabular language (frequency tables and contingency) are added.

- *Procedures:* We identified different procedures to compute frequencies and probabilities, based on the table data; procedures to perform the homogeneity and independence tests; computation of association measures (for 2x2 and rxc tables; based on the Chi-square statistics or reduction of proportional error), computing, reading and interpreting the chi-square table, and managing the software.
- *Concepts:* Among other concepts we highlight those of variable, value, frequency and its types; joint, marginal and conditional distribution, simple, compound and conditional probability; favourable and possible cases, sample / cell and population, percentage, ratio, proportion, independence / association, direct and inverse association, strength of association, hypothesis test: null and alternative hypotheses, region of acceptance / rejection, significance level p -value, chi-square, chi-square distribution, degrees of freedom.
- *Properties:* We analyzed the relationships between different types of frequencies, the properties of the frequencies in case of independence, the probability properties and axioms, the properties of the Chi-square statistic and distribution, degrees of freedom; and the meaning and variation of values in association measures.
- *Arguments:* We considered algebraic and/or deductive formal proofs, informal demonstrations, and checking examples and counterexamples, with no intention to generalize.

Although the review is limited to an elementary study of contingency tables (that could be expanded, for example, to the study of other association coefficients or prediction models based on these tables), the complexity of this mathematical object, simple in appearance, is visible. There is, however, a very limited teaching time, both in psychology (where we focused), as in other degrees. In addition, some procedures related to contingency tables – e.g. hypothesis tests- are themselves complex objects that were analyzed in previous research (Falk, 1986; Vallecillos, 1996; Lecoutre, Lecoutre & Pointevineau, 2001). All these indicators support the interest of a specific research focusing on the teaching of contingency tables.

These results were used to define the reference meaning of this work, to plan the teaching of the subject (Chapter 4) and to develop the teaching materials. It was also used to design the evaluation instruments (Chapters 5 and 6), where we tried to get

representativeness of the meaning elements to make the assessment more meaningful.

O2: Producing a research survey of the topic “contingency tables”. We started from the review included in Estepa’s (1993) thesis and took into account the subsequent work on the subject. We also tried to extend the synthesis carried out by this author.

To meet this objective, in Chapter 2 we present a synthesis of previous research related to contingency tables. This Chapter is organized into different sections related to conditioning and causation, the elements of causal relationship; philosophical theories of causation, meaning of causality from the statistical point of view, evolutionary development of association; strategies in association judgements; conceptions of association; evolution after teaching, and computing probabilities from contingency tables.

The analysis of previous research suggests that most of it was conducted within the field of psychology, beginning with the pioneering work by Piaget and Inhelder (1955) on the cognitive development of the concept and followed by others such as Smedlund (1963) Jenkins and Ward (1965), Shaklee and Tucker (1980), Allan and Jenkins (1983), Shaklee and Mins (1982) or Perez Echeverria (1990), who described adult’s strategies in judging association. Other psychologists studied the effect of previous theories, i.e., when subjects use their own beliefs, rather than statistical data when making an association judgment (Chapman & Chapman, 1969; Jennings, Amabile & Ross, 1982; Alloy & Tabacnick, 1984; Wright & Murphy, 1984).

The mathematics education approach, particularly by Estepa (1993) and colleagues; for example in Batanero, Estepa and Godino (1991, 1997), Estepa and Batanero (1995), Batanero, Estepa, Godino and Green (1996), Batanero, Godino and Estepa (1998) imply a paradigm shift. Their work is based on a much deeper mathematical and cognitive analysis, both of association tasks, as of students' responses to them. Through their analyses, they describe misconceptions and incorrect strategies; identify in-act-theorems implicit in the correct strategies and acts of understanding association produced along a teaching experiment based on the use of computers.

We also found research on computing probabilities in contingency tables that show the high relationship between the concepts of probability and association, as highlighted by Inhelder and Piaget (1955) and indicate the usefulness of contingency tables to work with marginal, conditional or joint probabilities or even more complex

theorems such as the of total probability or the Bayes theorem. These studies suggest problems in interpreting the associate probabilities, for example, in Maury (1985), Eddy (1982), Gotzsche and Olsen (2000), Martignon and Wassner (2002), Díaz (2007), Contreras, Estrada, Díaz and Batanero (2010), Krauss, Bruckmaier and Martignon, (2010), Contreras (2011), and Contreras, Batanero, Diaz and Fernandes (2011).

In summary, the state of art suggest a shortage of research on this subject, from the educational point of view, being an exception the work by Antonio Estepa and his colleagues. This finding reaffirmed our decision to investigate the teaching of contingency tables, a theme included in many university specialities, both in statistics and in other areas. Furthermore, it is common in these majors to include the teaching of homogeneity or association tests in contingency tables or the various measures of association. We therefore expect new difficulties as regard the student's general understanding of hypothesis tests and other specifically related to understanding of the Chi-square distribution, which had not been analyzed from the educational point of view.

O3: Conducting a study to evaluate the intuitive strategies, association judgments and estimation of the strength of association in contingency tables by psychology students before they formally study the subject.

To meet this goal, in Chapter 3 we present this study with a sample of 414 students from three Spanish universities. Firstly, we describe the process involved in adapting a questionnaire built by Estepa (1993) to our sample. We selected and adapted four items, taking also into account problems of contingency tables included in psychology textbooks. We analyzed the content for each item selected that were part of a questionnaire directed to assess the students' intuitive knowledge. We also described the task variables (Kilpatrick, 1978) taken into account in preparing the questionnaire: type of statistical dependence between the variables; strength of dependence, agreement between the data and the previous theories suggested by the problem context; and type of covariation.

The next step in achieving our goal was the pilot trial of the questionnaire, after which, some improvements were carried out, and the usefulness of the questionnaire for the research goals was confirmed (results of the pilot trial were published in Cañadas, Batanero, Contreras & Arteaga, 2011; 2012). The final questionnaire of intuitive

knowledge was given to a sample of 414 students; the results about four points: judgments of association, estimation of the association coefficient, previous theories about the association and analysis of strategies were analysed.

Regarding the association judgments, these results reflect that a majority of students considered the existence of association in all items, most often where there is a relationship in the data, but also in case of perfect independence, possibly due to illusory correlation (Chapman & Chapman, 1967, 1969; Jennings, Amabile & Ross, 1982; Wright & Murphy, 1984; Alloy & Tabacnick, 1984; Murphy & Medin, 1985).

When estimating the association coefficient, students generally always inferred a meaningful association between the variables. The difference with the coefficient true value was higher in case of perfect independence, possibly because of the effect of previous theories and because this item had the highest frequency in cell *a* of the table. In case of direct dependence, the estimation of the association coefficient was almost perfect; moreover that in this item the students did not show previous theories; the approximation for the inverse dependence item was not so good. In general, in the items where students had previous theories, they over-estimated the coefficient value, while they provided an almost perfect estimation when there were no theories, which agree with previous research (Chapman & Chapman, 1967, 1969; Jennings, Amabile & Ross, 1982; Wright & Murphy, 1984; Alloy & Tabacnick 1984; Medin & Murphy, 1985). Two types of strategies analyses were carried out: firstly, we studied the number of cells and type of frequency used, as well as the processing levels defined by Pérez Echeverría (1990); then we conducted a semiotic analysis of strategies to identify some semiotic conflicts (Godino, Batanero & Font, 2007) that could explain the students' failures.

The most common incorrect strategies for the different levels of Pérez Echeverría (1990), were the following: (a) using only the maximum frequency cell (SI.1.1) in level 1, a strategy also found by Smedlund (1963), Shaklee and Mins (1982) and Estepa (1993); (b) using only a conditional distribution (SI.2.1) in level 2; (c) comparing cell *a* (presence of both characters) with cells *b* and *c* (presence of only one character) (SI.3.1) in level 3; (d) using two conditional distributions (SP.4.2) in level 4; and (d) comparing the join relative frequencies (SI.5.3) in level 5. These strategies repeat, with some variations those described by Estepa (1993), so that our results indicate that in entering university, psychology students do not have enough competence to interpret and judge association in a contingency table.

We made an original classification of the semiotic conflicts identified, according whether they relate to association, independence, or other concepts and properties. We found those defined by Estepa (1993) as local and causal conception of association, which we prefer to describe as conflicts, whose meaning is weaker than that of conception, since we do not need to check its resistance. We also identified other conflicts not described by this author, some of which were found in other statistical topics (for example in relation to conditional probability, by Falk, 1996, Diaz, 2007; Contreras, 2011); however, most of them had not been identified in previous research and constitute an original contribution.

O4: Designing and experimenting a study process of contingency tables designed for psychology students, which addresses the major problems identified in Chapter 1, or in previous research (Chapter 2), those identified in the evaluation study (Chapter 3), and the curricular restrictions.

This goal is addressed in Chapter 4, where we discuss a study process consisting in four theoretical sessions in the traditional classroom and two practical sessions. We describe and analyze the material provided to the student, which consists of written materials, exercises, practical activities and an Excel program that facilitates the calculus. We also prepared a website to make this material available to the students. As a result of this analysis, we describe the institutional intended meaning in teaching. We also prepared assessment tools to evaluate learning.

At the end of Chapter 4, we also analyze the potential epistemic suitability of the study process, which was reasonable, since although the intended institutional meaning was more limited than the reference meaning, the proposal includes the most important elements of meaning of contingency tables and was adapted to the time available and the students' prior knowledge (cognitive suitability).

O5: Assessing the learning achieved by the students after the study process. This goal is addressed in Chapters 5 and 6, where we describe and analyze the assessment instruments, and discuss the assessed meaning. We also analyze the students' responses, describing the elements of meaning acquired and the semiotic conflicts remaining after the instruction.

In Chapter 5 we present the results of the assessment with questionnaires composed of true/false and multiple choice items that were built following the usual psychometric recommendations (Osterlind, 1989; Thorndike, 1989; Martinez Arias, 1995). Part of these results were published in Cañadas, Arteaga, Batanero and Contreras (2010) and Cañadas, Ortiz, Contreras and Gea (2012). The a priori analysis of the questionnaire suggests most of the elements of meanings included in the teaching experiment were taken into account, ensuring the content validity of the assessment process. Furthermore, the items difficulty and discrimination and appropriate values of reliability, imply the usefulness of the questionnaire as an assessment tool (Martinez, Hernandez & Hernandez, 2006). The average number of correct answers per student in the questionnaire was higher than the expected theoretical mean value, showing a quite satisfactory learning.

We performed a detailed analysis of the responses for each item. Among other results we highlight the following: some students did not reach the data extraction level (Bertin, 1967) in readings the tables; the students seemed to reason at level 5 of Perez Echeverria (1990), overcoming the additive or lower levels strategies observed in the initial assessment (Chapter 3); the students learnt the properties of independence and applied them in their association judgments, for example, to deduce the direct or inverse dependence; the students seemed to overcome the causal conception of the association described by Estepa (1993); recognized that in random dependence, to a value of the independent variable may correspond several of the dependent variable; learnt to compute the degrees of freedom, in contrast to Olivo (2008) and Alvarado (2007) research; recognized the assumptions needed to application the chi-square, and correctly interpreted Phi value, as well as the V and C association coefficients

Moreover, the item difficulty was varied, and we found a low proportion of correct answers on items related to the goal of building a contingency table, and the formula of conditional relative frequencies; it was also difficult to recognize that the odds ratio can take values over one, and the interpretation of the Chi-square and p values; as well as the relationship between the relative risk by rows and columns, and the interpretation of Lambda. Results on the final test were worse than expected, possibly because of the penalty students received in case of wrong answer or the effect of forgetting; correct answers were lower than 50% in all items.

Regarding the open-ended problems (Chapter 6), we notice the following results: students were able to use at least one correct method to study association,

having overcome their initial incorrect strategies and almost entirely reasoning at level 5 in Perez Echeverria's categorization (1990), which means that they used all the cells in the contingency tables with multiplicative strategies. Students disregarded the effect of previous theories, visible in the initial study, in their association judgments; instead they computed some formal some coefficient to estimate association. They operated with relative frequency, which indicates a better understanding of the idea of distribution; students quickly learnt the software management and efficiently computed conditional probabilities from contingency tables improving the results obtained in previous research (Diaz, 2007; Estrada & Díaz, 2007; Contreras, 2011). On the downside, they had difficulty in contextualizing their solutions.

Finally, we emphasize the quantitative results: the average number of correct responses in the whole set of problems (Chapter 6) was higher than the expected theoretical mean value, showing a pretty good learning.

7.3. CONCLUSIONS REGARDING THE HYPOTHESES

After discussion the findings in relation to the objectives, we next present those related to the research hypotheses.

Hypothesis 1. Students were expected to be influenced by their previous that would significantly affect their association judgments and their estimation of association.

This hypothesis was confirmed in the initial study (Chapter 3), where some psychology students were influenced by their previous theories, in agreement with previous research by Jennings, Amabile, and Ross (1982), Wright and Murphy (1984), Alloy and Tabacnick (1984), Perez Echeverria (1990), Estepa (1993) and Meiser and Hewstone (2006). This happened in all the items, where the number of students showing this influence varied between 24 (5.8%) and 40 (9.7%), especially in the independence item, where data contradicted their previous theories. The influence of these theories was also shown in the estimation of the association, as the best precision was obtained in items where there were no previous theories.

By contrast, this hypothesis was contradicted, after teaching, since the effect of previous theories was not observed in the open problems (Chapter 6). Consequently, the teaching seemed to help removing the intuitive use of previous theories that constantly

appears in association research in subjects without prior instruction. However, more research is needed to confirm this point, since after instruction we did not propose items where data clearly contradicted the students' theories.

Hypothesis 2. In the initial study students approach the association judgement with intuitive strategies, most of which are incorrect or partially correct. The percentage of level 5 strategies in Perez Echeverria's (1990) classification is small.

This hypothesis was also confirmed in the initial study (Chapter 3), where comparisons made among the three universities in the sample (Almería, Granada and Huelva) suggested the sample homogeneity and results similar to those obtained by Estepa (1993). A summary of strategies showed a majority incorrect strategies in 2x2 tables (between 59.4% and 64.6%) and in the whole questionnaire.

In Chapter 3 we also classified the strategies according to the Perez Echeverria's levels (1990). The most common strategy was using only a conditional distribution (level 2), except in item 4 where the use of all cells with additive comparisons was higher (level 4). The second more frequent level was level 4 (in 2x2 tables) and 2 (in the 2x3 table). Less than 20% of students achieved level 5, although this percent was higher than that indicated by Pérez Echeverría.

Hypothesis 3. The task variables and strategy used in the initial influence the association judgment, and the estimation of association strength.

In Chapter 3 we analyzed the relationships of the task variables considered in the items: type of statistical dependence; strength of association, correspondence between the data and the previous theories suggested by context and type of covariation, in Barbancho's classification (1973). The effect of previous theories was higher in the independence item, where students showed a causal conception of the association, linking the concepts of association and causation. In the item corresponding to inverse dependence we scarcely noted the unidirectional conception of association defined by Estepa (1993). On the other hand, results in the direct dependence item were very good, and similar to those in Estepa's study; 93.2% of students judged direct association while in Estepa the percent was 91.1%. Finally in the item corresponding to a 2x3 table most students provided a correct response (97.1%), with better results than those in Estepa's study (87.3%).

Hypothesis 4. We expected to identify different semiotic conflicts that may explain the incorrect or partially correct strategies in association judgments, as a part of the semiotic analysis of students' strategies in the initial study.

In Chapter 3, we performed a semiotic analysis and identified a number of semiotic conflicts, i.e., students' interpretations that do not correspond to what is expected by the teacher. We classified these conflicts into four separate blocks: (a) incorrect properties attributed to association; (b) incorrect properties attributed to independence; (c) confusion between concepts and properties; and (d) other conflicts. We found a wide variety of conflicts affecting the interpretation of contingency tables that teachers could identify and prevent in the future teaching of the topic. This includes the local and causal conceptions of association described by Estepa (1993) in a large part of the sample, which, as stated, we prefer describing as conflicts, which involves only confusion on the part of the student, more than a strong conviction. New conflicts not described by this author also appeared.

Hypothesis 5. We expected the acquisition of the most significance elements of meaning included in the study process by more than 50% of students taking part in the teaching experiment.

In Chapter 5, we provide a detailed analysis of the learning achieved by these students, which was assessed with questionnaires. We studied the items difficulty, discrimination and the questionnaire reliability, obtaining good results in all the three aspects. The data analysis showed that half the students answered more than 75% of questions correctly, which indicates very good learning outcomes.

Hypothesis 6. Some initial misconceptions described by Estepa (1993) would be overcome by the participating students after teaching.

Some skills that proved to be problematic in previous research were overcome by participants in our study after teaching. For example the causal conception of association, that was frequent in the work of Estepa (1993). Moreover, only few students were influenced by their previous theories after teaching; students also worked with relative frequencies and most of them used all the data, overcoming the local conception of association.

Hypothesis 7. We expected more correct strategies in association judgment after teaching and acquisition of new formal strategies after instruction, with small presence of strategies of level 4 or lower in Pérez Echeverría's classification.

After instruction, students improved their strategies and solution of problems. In Chapter 6, we analyzed some open-ended problems, showing the overcoming of some misconceptions, such as the local conception. Moreover, most students changed to level 5 multiplicative strategies (Pérez Echeverría, 1990). In addition, students used other more formal strategies, valid to obtain correct results; for example, the calculation of association coefficients. This led to the emergence of other difficulties such as not remembering the formulas or errors in hypothesis testing that coincided with some difficulties described by Vallecillos (1997) and Batanero (2000).

Hypothesis 8. Students were also expected to show competence in conducting homogeneity tests and interpreting the different association coefficients in open-ended problems.

The complexity of the hypothesis tests was described in a wide research (e.g. Vallecillos, 1996; Vallecillos & Batanero, 1997), most of which focused on parametric tests. Some of these difficulties might appear in the independence and homogeneity tests from contingency tables.

In Chapter 6, we examined an open-ended problem of homogeneity test, where 79.3% students posed correct or partially correct hypotheses (59.7% correctly posed the null hypothesis and 46.6% the alternative hypothesis). 63.1% of students correctly determined the statistics and p value; the Chi-square statistic was correctly computed by all the students, the degrees of freedom by 70.7% of the sample and the p -value by 63.1%, which gives us an idea of the relative difficulty of the three tasks. On the other hand, 51.9% students made the right decision (accepting or rejecting the null hypothesis) and 43.5% correctly interpreted the results in the problem context, in the last section, covering thus all the steps of the modelling process (Henry, 1997): presentation of hypotheses, defining and working with a mathematical model and interpreting the results.

Students were able to abstract and simplify reality, going from to the idea of emotional component, to the data distribution of 100 participants's perception; that is, accepting the problem data as indicators of the emotional component. They worked with mathematical models (random variables and statistics corresponding to the emotional

component perception, the Chi-square statistic, its degrees of freedom, p -value and significance level, logic of hypothesis testing, etc.). Finally the last step (interpreting the mathematical model results (decision made) in the problem context (translating this decision to what indicates about the emotional component of words) was the hardest part, in agreement with other previous studies (for example, Arteaga, 2011).

Hypothesis 9. We expected an adequate didactic suitability in the designed study process.

Finally we justified some indicators of the *didactic suitability* taken from our theoretical framework (Godino, 2011) to analyze the study process described in Chapters 4-6 in its different components (epistemic, cognitive, affective, interactional, meditational suitability).

- Firstly, the intended institutional meaning (analyzed in Chapter 4) showed a reasonable *epistemic suitability*, since we included in the study process the most important elements of the intended meaning for contingency tables as described in Chapter 1.
- The *cognitive suitability* was adequate, since the teaching was adapted to the available teaching time and the students' prior knowledge. We also made an effort to prepare a suitable material, including a spreadsheet program to facilitate learning. The results of assessment (Chapters 5 and 6) suggested that there was a good learning, and most planned objectives were met.
- Various teaching resources (time, students' organization, written materials, technology) were used throughout the study process that contributes to its *media suitability*.
- Finally, since the problem fields were contextualized in applications in the area of psychology to motivate the students, the proposal has an *affective and ecological suitability*, since it took into account the constraints and relations with other subjects. Of course these components are related, since, for example, the higher is the students' motivation the better learning results would be obtained (Gundlach, Kuntze, Engel y Martignon, 2010).

7.4. STUDY CONTRIBUTIONS AND LIMITATIONS

The research provided a number of contributions of interest for the teaching of contingency tables and the training of psychology professionals, the main of which are summarized below.

Firstly, we provided detailed information about the students' biases and reasoning difficulties related to contingency tables, both in the analysis of the institutional referential meaning (Chapter 1) as well as in the synthesis of previous research (Chapter 2).

We developed and adapted questionnaires that may be useful for the study of contingency tables in the field of psychology, basing on previous research (Estepa, 1993) and statistics books oriented to psychology or education (e.g. Ruiz-Maya, Martin, Montero & Uriz, 1995; Ato & Lopez, 1996; Lipschutz & Schiller, 1999; Ruiz-Maya & Martin, 2002; Batanero & Diaz, 2008). These questionnaires are useful for both the study of intuitive knowledge of contingency tables (Chapter 3) and for the formal study of the subject (Chapters 5 and 6).

In the work we have developed materials to assist teachers in the classroom that could be used in traditional classrooms or in teaching based on technology and were made available on Internet (www.ugr.es/~analisisdedatos/webcurso/presentacion.html) . These materials (besides the questionnaires already mentioned) include presentations, a booklet (Cañadas, 2011), exercises, the Excel program, and the program description.

We found no previous teaching experiments of the topic in psychology, so our study in this field is new. The initial study of intuitive knowledge (Chapter 3) where most biases found by other researchers still appeared showed the need to further explore this issue in psychology training. The wide sample size (415 students) strength the results; at the same time, new biases appeared to serve as reference in the future.

We planned and developed a teaching process (Chapter 4) that sought to eliminate the biases found in previous research and was based on all the materials described above. The detailed results of the assessment of this experiment (with 94 students) is also an important information for future planning of instruction.

Limitations of the study

Like any empirical study, this research has some limitations. The samples used in both the assessment of intuitive knowledge and in the teaching experiment were non random, although wide (414 participants in the first study and 94 in the second).

Moreover, the geographical context was limited to three provinces of Andalucía (Almería, Granada and Huelva). It would then be useful to collect more data from other contexts to confirm our results.

The teaching materials developed, could be improved by taking into account the results from the student's assessment in Chapters 5 and 6; in particular deepening those points that were harder. As regard to the teaching process developed, the main limitation was the time available, which had to be taken into account in the selection of content to be taught. This content could be expanded in future research with a longer teaching period.

The information was mainly collected by questionnaires (except the teaching observation). This means no direct interaction with students during the information collection. These results could be complemented by in-depth interviews to students to better analyse their reasoning and their incorrect responses.

Finally, the work focused primarily on the teaching of contingency tables for psychology students. We could vary the type of students and adapt the materials and questionnaires if to be used in other fields, such as medicine.

7.5. IDEAS FOR FUTURE RESEARCH

Because of the above limitations in our work, new lines to continue the research on contingency tables arise.

We could alter the task variables analyzed in the initial study to check whether this affect our results. For example, we could analyze the sign of association (which was not considered in our work), or add 3x3 contingency tables. In addition, as mentioned above, we could collect information outside of Andalucía, to see if the geographical context affects the results. On the other hand, in the items selected the students were given the contingency tables; it could be of interest to consider a problem where they had to build the table. Furthermore we have not studied the construction of graphs, despite being included in teaching, and a potential source of information (Arteaga, 2011).

Besides adapting the materials developed to other university major, such as medicine or sociology it would also be possible to make a comparative study of teaching, in these new contexts, as well as at high school level.

In summary, as a synthesis of this document, we think we reported enough research findings and implications for instruction, and provided new teaching materials and assessment instrument that may be also useful in further research. All of this fulfils the standard required to obtain a doctorate degree.

REFERENCIAS

- ANECA (Agencia Nacional de la Evaluación de la Calidad y Acreditación) (2005). *Libro blanco. Título de Grado en Psicología*. Madrid: Autor.
- ANECA (Agencia Nacional de la Evaluación de la Calidad y Acreditación) (2010). Documento de verificación del título de Grado en Psicología. Universidad de Granada. Online: vicengp.ugr.es/pages/_grados-verificados/51psicologiaverificado
- Aguilera, A. M. (2001). *Tablas de contingencia bidimensionales*. Madrid: La Muralla.
- Allan, L. G. y Jenkins H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation*, 14, 381-405.
- Alloy, L. B. y Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, 91, 112-149.
- Alvarado, H. (2007). *Significados del teorema central del límite en la enseñanza de la estadística en ingeniería*. Tesis Doctoral. Universidad de Granada.
- Álvarez, H. y Pérez, E. (2004). Causalidad en medicina. *Historia y filosofía de la medicina. Gaceta Médica de México*, 140(4), 467-472.
- Amón, J. (1993). *Estadística para Psicólogos 2*. Madrid: Pirámide.
- Anguera, M.T. (1990). Metodología observacional. En J. Arnau, M.T. Anguera y J. Gomez (Ed.), *Metodología de la investigación en ciencias del comportamiento* (pp. 123-236). Universidad de Murcia.
- Anguera, M. T. (1993a), *Metodología observacional en la investigación psicológica*, Vol I. Barcelona: PPU.
- Anguera, M. T. (1993b) *Metodología observacional en la investigación psicológica*, Vol II. Barcelona: PPU.
- Aoyama. K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education*, 2(3). On line: <http://www.iejme/>.
- Arkes, H. R. y Harkness, A. R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General*, 112 (1), 117-135.
- Arteaga, P. (2008). *Análisis de gráficos estadísticos elaborados en un proyecto de análisis de datos*. Trabajo fin de Master. Departamento de Didáctica de la Matemática.
- Arteaga. P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y*

- conocimientos didácticos de futuros profesores*. Tesis Doctoral. Universidad de Granada.
- Arteaga, P., Batanero, C., Cañadas, G. y Contreras, J. M. (2011). Las tablas y gráficos estadísticos como objetos culturales. *Números* 76, 55-67.
- Ato, M. y López, J. J. (1996). Análisis estadístico para datos categóricos. Madrid: Síntesis Psicología.
- Bakan, D. (1966). The tests of significance in psychological research. *Psychological Bulletin*, 66, 423-437.
- Ball, D. L., Lubienski, S. T. y Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. En V. Richardson (Ed.), *Handbook of research on teaching* (pp. 433-456). Washington, DC: American Educational Research Association.
- Barbancho, A. G. (1973). *Estadística elemental moderna*. Barcelona: Ariel. (Cuarta edición).
- Barbero, M. (2003). *Psicometría II. Métodos de elaboración de escalas*. Madrid: UNED.
- Batanero, C. (2000). Controversias around the role of statistical test in experimental research. *Mathematical Thinking and Learning*, 2 (1-2), 75-98.
- Batanero, C. (2001). *Didáctica de la estadística*. Departamento de Didáctica de la Matemática. Universidad de Granada.
- Batanero, C. (2002). Los retos de la cultura estadística. Conferencia en las *Jornadas Interamericanas de Enseñanza de la Estadística*. Buenos Aires. Confederación Latino-americana de Sociedades de Estadística.
- Batanero, C., Burrill, G., Reading, A y Rossman, A. (2008) (Eds.). *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. Monterrey: ICMI e IASE. On line: www.stat.auckland.ac.nz/~iase/publicatons.
- Batanero, C., Cañadas, C., Contreras, J. M. y Arteaga, P. (2012). É sinxela a interpretación das táboas de continxencia? *Gamma*, 12, 27-34.
- Batanero, C., Cañadas, C., Estepa, A. y Arteaga, P. (2012). Psychology students' estimation of association. En Tai-Yih Tso (Ed.), *Proceedings of 36th Conference of the International Group for Psychology in Mathematics Education* (Vol. 2, pag. 51-58). Taipei, Taiwan: PME.

- Batanero, C. y Díaz, C. (2005). Análisis del proceso de construcción de un cuestionario sobre probabilidad condicional. Reflexiones desde el marco de la TFS. En A Contreras (Ed.), *Investigación en Didáctica de las Matemáticas* (pp. 13-36). Granada: Grupo FQM126.
- Batanero, C. y Díaz, C. (2008). *Análisis de datos con Statgraphics*. Granada: Departamento de Didáctica de la Matemática.
- Batanero, C., Estepa, A. y Godino, J. D. (1991). Estrategias y argumentos en el estudio descriptivo de la asociación usando microordenadores. *Enseñanza de las Ciencias*, 9(2), 145-150.
- Batanero, C., Estepa, A. y Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer based teaching environment. En J. B. Garfield y G. Burrill (Ed.), *Research on the role of technology in teaching and learning statistics* (pp. 191-206). Minnesota, MN: International Statistical Institute.
- Batanero, C., Estepa, A., Godino, J. y Green D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, 27(2), 151-169.
- Batanero, C., Godino, J. D. y Estepa, A. (1998). Building the meaning of statistical association through data analysis activities. *Research Forum*. En A. Olivier y K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education* (v.1, pp. 221-236, con discusión en pp. 237-242). University of Stellenbosch.
- Batanero, C., Godino, J. D., Green, D. R., Holmes, P. y Vallecillos, A. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.
- Bertin, J. (1967). *Semiologie graphique*. Paris: Gauthier-Villars.
- Beyth-Marrom, R. (1982). Perception of correlation reexamined. *Memory and cognition*, 10 (6), 511-519.
- Birnbaum, I. (1982). Interpreting statistical significance. *Teaching Statistics*, 4, 24-27.
- Bolívar, A. (1998). Tiempo y contenido del discurso curricular en España. *Revista de currículum y formación del profesorado*, 2(2), On line: www.ugr.es/~recfpro/rev22ART4.pdf.
- Bollen, K. A. (1989). *Structural equations with latent variables*. Nueva York: Wiley.

- Borges, A., San Luis, C., Sánchez, J. A. y Cañadas, I. (2001). El juicio contra la hipótesis nula: muchos testigos y una sentencia virtuosa. *Psicothema*, 13 (1), 174-178.
- Bradford-Hill A. (1992). Ambiente y enfermedad: ¿asociación o causación? *Boletín de la oficina sanitaria panamericana*, 113, 233-242.
- Brewer, J. K. (1986). Behavioural statistics textbooks: source of myths and misconceptions? En R. Davidson y J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics*, (pp. 127-131). Victoria, Canada: University of Victoria.
- Brewer, J. K. (1989). Analogies and parables in the teaching of statistics. *Teaching Statistics*, 11, 21-23.
- Brousseau, G. (1997). *Theory of didactical situations in mathematics. Didactique des mathématiques, 1970-1990*. Dordrecht: Kluwer.
- Bruno, A. y Espinel, M. C. (2009). Construction and evaluation of histograms in teacher training. *International Journal of Mathematical Education in Science and Technology*, 40(4), 473-493.
- Cañadas, G. R. (2011). *Las tablas de contingencia para psicología*. Granada: Departamento de Didáctica de la Matemática.
- Cañadas, G. R., Arteaga, P. y Batanero, C. (2010). Las tablas de contingencia: un objeto semiótico complejo. Trabajo presentado en el *XIV Simposio de la Sociedad Española de investigación en Educación Matemática*. Badajoz: Sociedad Española de Investigación en Educación Matemática.
- Cañadas, G. R., Arteaga, P., Batanero, C. y Contreras, J.M. (2010). Comprensión de tablas de contingencia e implicaciones didácticas. En A. Pérez (Ed.), *Actas del XIII Congreso de Enseñanza y Aprendizaje de las Matemáticas* (pp. 186-191). Córdoba: Sociedad Thales de Educación Matemática.
- Cañadas, G. R., Batanero, C., Contreras, J. M. y Arteaga, P. (2011). Estrategias en el estudio de la asociación en tablas de contingencia por estudiantes de psicología. *Educacion Matemática*, 23(2), 5-32.
- Cañadas, G. R., Batanero, C., Contreras, J. M. y Arteaga, P. (2012). Estimación de la asociación en tablas de contingencia por estudiantes de psicología. *UNO*, 60, 87-94.
- Cañadas, G. R., Batanero, C., Díaz, C. y Gea, M. M. (2012). Comprensión del test chi-cuadrado por estudiantes de psicología. En A. Estepa, A. Contreras, J. Deulofeu,

- C. Penalva, F. García y L. Ordóñez (Ed.), *Actas del XVI Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 153-163). Baeza: Sociedad Española de Investigación en Educación Matemática.
- Cañadas, G. R., Batanero, C., Díaz, C., Roa, R. (2012) Psychology students' understanding of the chi-squared test. *Statistique et Enseignement*, 3(1), 3-18.
- Cañadas, G. R., Díaz, C., Batanero, C. y Arteaga, P. (2011). Estimación de la asociación por estudiantes de psicología. En M. Marín, G. Fernández, L. Blanco y M. Palarea (Eds.), *Investigación en Educación Matemática XV Simposio de la Sociedad Española de Educación Matemática* (pp. 297-304). Ciudad Real: Sociedad Española de Investigación en Educación Matemática.
- Cañadas, G. R., Ortiz, J. J., Contreras, J. M. y Gea M. M. (2012). Enseñanza y comprensión formal de las tablas de contingencia. Trabajo presentado en la 26 *Reunión Latinoamericana de Matemática Educativa*. Ouro Prieto, Brasil.
- Carmine, E. G. y Zeller, R. A. (1979). *Reliability and validity assessment*. London: Sage University Paper.
- Carrión, J. C., & Espinel, M. C. (2006). An investigation about translation and interpretation of statistical graphs and tables by students of primary education. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador, Bahia, Brazil: International Statistical Institute and International Association for Statistical Education. Online: www.stat.auckland.ac.nz/~iase/publications.
- Carvalho, C. (2001). *Interação entre pares. Contributos para a promoção do desenvolvimento lógico e do desempenho estatístico no 7º ano de escolaridade*. Tesis Doctoral. Universidad de Lisboa.
- Castro-Posada, J. (2001). *Metodología de la investigación. Fundamentos*. Salamanca: Sage.
- Castro Sotos, A. E., Vanhoof, S., Van den Norrgate, W. y 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.
- Cazorla, I. (2002). *A relação entre a habilidades viso-pictóricas e o domínio de conceitos estatísticos na leitura de gráficos*. Tesis Doctoral. Universidad de Campinas.
- Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal*

Learning and Verbal Behavior, 6(1), 151-155

- Chapman, L. J. y Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 72, 193-204.
- Chapman, L. J. y Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs, *Journal of Abnormal Psychology*, 74, 271-280.
- Chow, L. S. (1996). *Statistical significance: Rational, validity and utility*. London: Sage.
- Clark, P. (1994). ¿Cómo se establece la causalidad en medicina? En H. M. Ramiro y O. Saita-Kamino (Eds.), *Temas de medicina interna. Epidemiología clínica* (pp. 265-73). México: Interamericana-McGraw-Hill.
- Cobb, P. y Mc Clain, K. (2004). Principles of instruccional design for supporting the development of students statistical reasoning. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking*. Netherlands: Kluwer Academic Publishers.
- Cobo, B. (2003). *Significados de las medidas de posición central para los estudiantes de secundaria*. Tesis Doctoral. Universidad de Granada.
- Cockcroft, W. H. (1985). *Las matemáticas sí cuentan*. Madrid: M.E.C.
- Contreras, J. M. (2011). *Evaluación de conocimientos y recursos didácticos en la formación de profesores sobre probabilidad condicional*. Tesis Doctoral. Universidad de Granada.
- Contreras J. M., Batanero, C. Díaz, C. y Fernandes, J. A. (2011). Prospective teachers' common and specialized knowledge in a probability task. En M. Pytlak, T. Rowland y E. Swoboda (Eds.), *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education*. Rzeszow, Polonia: ERME. Online: sdoukakis.files.wordpress.com/2012/01/sd_mxm_dz_cerme7.pdf
- 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: Sociedad Española de Investigación en Educación Matemática.
- Cook, T. D. y Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis issues for field setting*. Chicago: Rand Mc. Nelly Publishing Company.

- Cuadras, C. M., Echeverría, B., Mateo, J. y Sánchez, P. (1984). *Fundamentos de Estadística Aplicación a las Ciencias Humanas*, Promociones Publicaciones Universitarias, Barcelona.
- Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: N.C.T.M.
- D'Amore, B. (2000). La complessità dell'educazione e della costruzione dei saperi. *Riforma e didattica*, 4, 35-40.
- Danhke, G. L. (1989). Investigación y comunicación. En C. Fernández-Collado y G. L. Danhke (Eds.), *La comunicación humana: ciencia social* (pp. 385-454). México, D. F.: McGraw-Hill.
- Davis, J. A. (1985). *The logic of causal order*. Londres: Sage University Paper.
- Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología*. Tesis doctoral. Universidad de Granada.
- Díaz, C., Batanero, C. y Wilhelmi, M. (2008). Errores frecuentes en el análisis de datos en educación y psicología. *Publicaciones*, 35, 109-123.
- Díaz, C. y de la Fuente, E. I. (2005a). Conflictos semióticos en el cálculo de probabilidades a partir de tablas de doble entrada. *Biaix* 24, 85, 91.
- Díaz, C. y de la Fuente, E. I. (2005b). Razonamiento sobre probabilidad condicional e implicaciones para la enseñanza de la estadística. *Epsilon*, 59, 245-260.
- Díaz, J. y Gallego, B. (2006). Algunas medidas de utilidad en el diagnóstico. *Revista Cubana de Medicina General Integrada*, 22 (1).
- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. En D. Kahneman, P. Slovic y Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases*. Nex York: Cambridge University Press.
- Einhorn, H. J. y Hogart, R. M. (1986). Judging probable cause. *Psychological Bulletin*. 99, 3-19.
- Ellet, F. S. y Erickson, D. P. (1986). Correlation, partial correlation and causation. *Synthese*, 67, 157-173.
- Engel, J. y Sedlmeier, P. (2011). Correlation and regression in the training of teachers. En C. Batanero, G. Burrill, and C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education: A Joint ICMI/IASE study* (pp. 247-258). New York: Springer.
- Erlick, D. E. y Mills, R. G. (1967). Perceptual quantification of condicional dependency, *Journal of Experimental Psychology*, 73 (1), 9-14.
- Espinel, C. (2007). Construcción y razonamiento de gráficos estadísticos en la

- formación de profesores. *Investigación en Educación Matemática 11*, 99-119.
- Estepa, A. (1993). *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.
- Estepa, A. (2004). Investigación en Educación Estadística. La asociación estadística. En R. Luengo (Ed.). *Líneas de investigación en Educación Matemática*, (pp. 227-255). Badajoz: Servicio de Publicaciones. Universidad de Extremadura.
- Estepa, A. y Batanero, C. (1995). Concepciones iniciales sobre la asociación estadística. *Enseñanza de las Ciencias*, 13(2), 155-170.
- Estrada, A. y Díaz, C. (2006). Computing probabilities from two way tables. An exploratory study with future teachers. En A. Rossman y B. Chance (Eds.), *Proceedings of Seventh International Conference on Teaching of Statistics*. Salvador (Bahia): International Association for Statistical Education. CD ROM.
- Estrada, A. y Díaz, A. (2007). Errores en el cálculo de probabilidades en tablas de doble entrada en profesores en formación. *UNO*, 44, 48-58.
- Euler, L. (1770). *Vollständige Anleitung zur Algebra. Vol. I*. San Petersburgo, Rusia: Kays. Academia der Wissenschaften.
- Falk, R. (1986). Conditional probabilities: insights and difficulties. En R. Davidson y J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics*. (pp. 292-297). Victoria, Canada: International Statistical Institute.
- Fletcher, R. H., Fletcher, S. W. y Wagner, E. H. (1998). *Epidemiología clínica. Aspectos fundamentales*. 2ª ed. México: Masson Williams.
- Font, V. y Godino, J. D. (2006). La noción de configuración epistémica como herramienta de análisis de textos matemáticos: su uso en la formación de profesores. *Educação Matemática Pesquisa*, 8 (1), 67-98.
- Font, J. D., Godino, J. D. y D'Amore, B. (2007). An ontosemiotic approach to representations in mathematics education. *For the Learning of Mathematics*, 27 (2), 3-9.
- Font, V., Godino, J. D. y Gallardo, J. (2012). The emergence of objects from mathematical practices. *Educational Studies in Mathematics*. En prensa.
- Font, V., Planas, N. y Godino, J. D. (2010). Modelo para el análisis didáctico en educación matemática. *Infancia y Aprendizaje*, 33 (2), 89-105.

- Franklin, C. y Mewborn, D. (2006). The statistical education of PreK-12 teachers: A shared responsibility. En G. Burrill (Ed.), *NCTM 2006 Yearbook: thinking and reasoning with data and chance* (pp. 335-344). Reston, VA: NCTM.
- Frías, M. D., Pascual, J. y García, J. F. (2000). Tamaño del efecto del tratamiento y significación estadística. *Psicothema*, 12(2), 236-240.
- Friel, S., Curcio, F. y Bright, G. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education* 32(2), 124-158.
- Gal, I. (2002). Adult's statistical literacy: Meaning, components, responsibilities. *International Statistical Review* 70(1), 1-25.
- Garfield, J. y Ben-Zvi (2008). *Developing students' statistical reasoning*. New York: Springer.
- Ghiglione, R. y Matalón, B. (1991). *Les enquêtes sociologiques. Théorie et pratique* París: Armand Colin.
- Gigerenzer, G. (1993). The superego, the ego and the id in statistical reasoning. En: G. Karen y C. Lewis (Eds.), *A handbook for data analysis in the behavioural sciences: Methodological issues* (p. 311-339). Hillsdale, NJ: Lawrence Erlbaum.
- Godino, J. D. (1996). Mathematical concepts, their meanings, and understanding. En L. Puig and A. Gutierrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (pp. 2-417-424), University of Valencia.
- Godino, J. D. (2002a). Perspectiva ontosemiótica de la competencia y comprensión matemática. *La matematica e la sua didattica*, 4, 434-450.
- Godino, J. D. (2002b). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, 22 (2 y 3), 237-284.
- Godino, J. D. (2011). Indicadores de la idoneidad didáctica de procesos de enseñanza y aprendizaje de las matemáticas. Conferencia en el *XIII CIAEM-IACME*, Recife, Brasil
- Godino, J. D. (2012). Origen y aportaciones de la perspectiva ontosemiótica de investigación en Didáctica de la Matemática. En A. Estepa, A. Contreras, J. Deulofeu, M. C. Penalva, F. J. García y L. Ordóñez (Eds.), *Investigación en Educación Matemática XVI* (pp. 49 - 68). Jaén: SEIEM.
- Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos

- matemáticos. *Recherches en Didactique des Mathématiques*, 14 (3), 325-355.
- Godino, J. D. y Batanero, C. (1998). Clarifying the meaning of mathematical objects as a priority area of research in Mathematics Education. En A. Sierpinska (Ed.), *Mathematics education as a research domain: A search for identity* (pp. 177-195). Dordrecht: Kluwer.
- 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.
- Godino, J. D., Batanero, C. y Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics*, 60 (1), 3-36.
- Godino, J. D., Bencomo, D., Font, V. y Wilhelmi, M. R. (2006). Análisis y valoración de la idoneidad didáctica de procesos de estudio de las matemáticas. *Paradigma*, XXVII (2), 221-252.
- Godino, J. D., Contreras, A. y Font, V. (2006). Análisis de procesos de instrucción basado en el enfoque ontológico-semiótico de la cognición matemática. *Recherches en Didactiques des Mathématiques*, 26 (1), 39-88.
- Godino, J. D., Font, V. y Wilhelmi, M. R. (2008). Análisis didáctico de procesos de estudio matemático basado en el enfoque ontosemiótico. *Publicaciones*, 38, 25-48.
- Godino, J. D., Font, V., Wilhelmi, M. R. y Castro, C. de (2009). Aproximación a la dimensión normativa en Didáctica de la Matemática desde un enfoque ontosemiótico. *Enseñanza de las Ciencias*, 27(1), 59-76.
- Godino, J. D., Wilhelmi, M. y Bencomo, D. (2005). Suitability criteria of a mathematical instruction process. A teaching experience of the function notion. *Mediterranean Journal for Research in Mathematics Education*, 4(2), 1-26.
- González, M. P. (2005). Probability and causality in the philosophy of Max Born. *LOGOS. Anales del Seminario de Metafísica* 38, 241-269.
- González, T., Espinel, C. & Ainley, J. (2011). Teachers' graphical competence. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. A Joint ICMI and IASE Study* (pp. 187-197). New York: Springer.
- Gotzsche, P. y Olsen, O. (2000). Is screening for breast cancer with mammography justifiable? *Lancet* 355, 129-34.

- Guàrdia, J., Freixa, M., Peró, M. y Turbany, J. (2007). *Análisis de datos en psicología*, Delta, Publicaciones Universitarias.
- Hagod, M. J. (1970). The notion of hypothetical universe. En D. E. Morrison y R. E. Henkel (Eds.), *The significance test controversy: A reader* (pp. 65-79). Chicago: Aldine.
- Haller, H. y Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research*, 7(1), 1–20.
- Harlow, L. L., Mulaik, S. A. y Steiger, J. H. (1997). *What if there were no significance tests?* Mahwah, NJ: Lawrence Erlbaum.
- 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. A Joint ICMI/IASE Study.* (pp. 235-246). New York: 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.
- Hernández, R., Fernández, C. y Baptista, P. (1994). *Metodología de la investigación*. McGraw-Hill.
- Huerta, M. P. y Lonjedo, M. A. (2003). La resolución de problemas de probabilidad condicional: un estudio exploratorio con estudiantes de bachiller. Comunicación presentada en el *Séptimo Simposio de la Sociedad Española de Investigación en Educación Matemática*. Granada: Sociedad Española de Investigación en Educación Matemática.
- Inhelder, B. y Piaget, J. (1955). *De la logique de l'enfant à la logique de l'adolescent*. París: 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.
- Jennings, D. L., Amabile, T. M. y Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgments. En D. Kahneman, P. Slovic y A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 211-230). Nueva York: Cambridge University Press.
- Kahneman, D., Slovic, P. y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Kilpatrick, J. (1978). Variables and methodologies in research on problem solving. En L. L. Hatfield y D. A. Bradboard (Eds.), *Mathematical problem solving: papers*

- from a research workshop* (pp. 7-20). Columbus: Eric/Smeac.
- Konold, C., Pollatsek, A., Well, A., Gagnon, A. (1997) Students analyzing data: Research of critical barriers. En J. Garfield. y G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 151-16). Voorburg, The Netherlands: International Statistical Institute.
- Krauss, S., Bruckmaier, G. y Martignon, L. (2010). Teaching young grown ups how to use Bayesian networks. En C. Reading (Ed.), *Proceedings of the Eight International Conference on Teaching Statistics*. Voorburg: IASE. Online: www.stat.auckland.ac.nz/~iase/publications.
- Krauss, S. y Wassner, C. (2002). How significance test should be presented to avoid the typical misinterpretations. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town: International Statistical Institute and International Association for Statistical Education. Online: www.stat.auckland.ac.nz/~iase/publications.
- Krippendorff, K. (1991). *Metodología de análisis de contenido*, Barcelona, Paidós.
- Laborda, E. (2007). Estudio del nexo de causalidad. *Cuadernos de Valoración*, 6, 37-47
- Laskey, K., y Martignon, L. (2010). Exploration and induction versus confirmation and deduction. En C. Reading (Ed.), *Proceedings of the Eight International Conference on Teaching Statistics*. Voorburg: IASE. Online: www.stat.auckland.ac.nz/~iase/publications.
- Lecoutre, B. (1999). Beyond the significance test controversy: Prime time for Bayes? *Bulletin of the Internacional Statistical Institute: Proceedings of the Fifty-second Session of the Internacional Statistical Institute* (Tome 58, Book 2) (p. 205-208). Helsinki: Internacional Statistical Institute.
- Lecoutre, B. (2006). Training students and researchers in Bayesian methods for experimental data analysis. *Journal of Data Science*, 4, 207-232.
- Lecoutre, B., Lecoutre, M. P., y Pointevineau, J. (2001). Uses, abuses and misuses of significance tests in the scientific community: Won't the Bayesian choice be unavoidable? *International Statistical Review*, 69, 399-418.
- León, O. G. y Montero, I (2002). *Métodos de investigación en psicología y educación*. Madrid: McGraw-Hill.
- Lipschutz, S. y Schiller, J. (1999). *Introducción a la probabilidad y estadística*. Madrid. Mc Graw Hill.
- Lipset, S. M., Trow, M. A. y Coleman, J. S. (1970). Statistical problems. En D. E.

- Morrison y R. E. Henkel, (Eds), *The significance tests controversy: A reader* (p. 81-86). Chicago: Aldine.
- Lonjedo, M. A. (2003). *La resolución de problemas de probabilidad condicional. Un estudio exploratorio con estudiantes de bachillerato*. Universidad de Valencia. Memoria de tercer ciclo.
- Lonjedo, M. A. y Huerta, M. P. (2005). La naturaleza de las cantidades presentes en el problema de probabilidad condicional. Su influencia en el proceso de resolución del problema. En A. Maz, B. Gómez y M. Torralbo (Eds.), *Investigación en Educación Matemática. Noveno Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 261-269). Universidad de Córdoba.
- Losada, J. L. (1999). *Metodología observacional*. A Coruña: Penta.
- Losada, J. L. y López-Feal, R. (2003). *Métodos de investigación en ciencias humanas y sociales*. Madrid: Thompson.
- Lubin, P., Maciá, M. A. y Rubio, P. (2005). *Psicología matemática II*, Volumen I, Universidad Nacional de Educación a Distancia.
- Martignon, L. y Wassner, K. (2002). Teaching decision making and statistical thinking with natural frequencies. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Voorburg: IASE. Online: <http://www.stat.auckland.ac.nz/~iase/publications/>.
- Martin, A. y Luna, J. D. (1994). *Bioestadística para las ciencias de la salud*, Ediciones Norma.
- Martínez, R. (1995). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid. Síntesis.
- Martínez, M. R., Hernández, M. J. y Hernández, M. V. (2006). *Psicometría*. Alianza Editorial.
- Maury, S. (1985). Influence de la question dans una épreuve relative á la notion d'independance. *Educational Studies in Mathematics*, 16, 283-301.
- Mayén, S. (2009). *Comprensión de las medidas de tendencia central por estudiantes mexicanos de Educación Secundaria y Bachillerato*. Tesis Doctoral. Universidad de Granada.
- Mayén, S., Díaz, C. y Batanero, C. (2009). Conflictos semióticos de estudiantes con el concepto de mediana. *Statistics Education Research Journal*, 8 (2), 74-93.
- McKenzie, C. y Mikkelsen, L. (2007). A bayesian view of covariation assessment. *Cognitive Psychology*, 54, 33-61.

- MEC (2006). *Real Decreto 1513/2006*, de 7 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación primaria.
- Meiser, T. y Hewstone, M. (2006). Illusory and spurious correlations: Distinct phenomena or joint outcomes of exemplar-based category learning? *European Journal of Social Psychology*, 36, 315–336.
- Meliá, J. L. (2001). *Teoría de la fiabilidad y la validez*. Valencia: Cristóbal Serrano.
- Menon, R. (1993). Statistical significance testing should be discontinued in mathematics education research. *Mathematics Education Research Journal*, 5 (1), 4-18.
- Messick, S. (1992). Validity. En R. L. Linn (Ed.), *Educational measurement* (pp. 13-104). Londres: MacMillan.
- Monteiro, C. y Ainley, J. (2004). Critical sense in interpretation of media graphs. En A. Cockburn y E. Nardi (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3. pp. 361-368). Norwich, UK: East Anglia University.
- Monteiro, C. y Ainley, J. (2007). Investigating the interpretation of media graphs among student teachers. *International Electronic Journal of Mathematics Education* 2 (3), 188-207. On line: <http://www.iejme/>.
- Moore, D. S. (1997). New pedagogy and new content. The case of statistics. *International Statistical Review*, 65(2), 123-165.
- Moreno, R. (1998). Causalidad en el modelo de campo: aportaciones metodológicas. *Acta Comportamentalia: Revista Latina de Análisis del Comportamiento*, 6, 59-71.
- Morrison, D. E. y Henkel, R. E. (Eds.). (1970). *The significance test controversy. A reader*. Chicago: Aldine.
- Murphy, G. L. y Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92 (3), 289-316.
- Nortes Checa, A. (1993). *Estadística teórica y aplicada*. Barcelona: PPU.
- Nisbett, R. y Ross, L. (1980). *Human inference: strategies and shortcomings of social judgment*. Nueva Jersey: Prentice Hall.
- Ojeda, A. M. (1995). Dificultades del alumno respecto a la probabilidad condicional. *UNO*, 5, 37-55.
- Olivo, E. (2008). *Significados del intervalo de confianza en la enseñanza de la ingeniería en México*. Tesis Doctoral. Universidad de Granada.
- Ortega Martínez, A. R. (1991). *Contingencia y juicios de covariación en humanos*.

Tesis Doctoral. Universidad de Granada.

- Osterlind, S. J. (1989). *Constructing test items*. Boston: Kluwer.
- Pardo, A. y San Martín, R. (1998). *Análisis de datos en psicología*, Pirámide.
- Peña, D. y Romo, J. (1997). *Introducción a la estadística para las ciencias sociales*. Madrid: McGraw-Hill.
- Perales, J., Catena, A., Ramos, M. y Maldonado, A. (1999). Aprendizaje de relaciones de contingencia y causalidad: Una aproximación a las tendencias teóricas actuales. *Psicológica*, 20, 163-193.
- Pérez Echeverría, M. P. (1990). *Psicología del razonamiento probabilístico*. Madrid: Ediciones de la Universidad Autónoma de Madrid.
- Piaget, J. (1975). *Psicología de la inteligencia*. Psique. Buenos Aires.
- Piaget, J. e Inhelder, B. (1951). *La gènèse de l'idée de hasard chez l'enfant*. Paris: PUF.
- Pollatsek, A., Well, A. D., Konold, C. y Hardiman, P. (1987). Understanding conditional probabilities. *Organisation, Behavior and Human Decision Processes*, 40, 255–269.
- Popper, K. (1962). *La lógica de la investigación científica*. Madrid: Tecnos.
- Postigo, Y. y Pozo, J. I. (2000). Cuando una gráfica vale más que 1000 datos: la interpretación de gráficas por alumnos adolescentes. *Infancia y Aprendizaje*, 90, 89 - 110.
- Pozo, J. I. (1987). *Aprendizaje de la ciencia y pensamiento causal*. Madrid: Síntesis.
- Retolaza, J. L. (2007). El problema de causalidad en las ciencias económicas. *Cuadernos de Gestión*, 7(2), 39-53.
- Ridgway, J., Nicholson, J. y McCusker, S. (2008). Mapping new statistical literacies and iliteracies. Trabajo presentado en el *11th International Congress on Mathematics Education*, Monterrey, Mexico.
- Ríos, S. (1976). *Análisis de decisiones*. Madrid: ICE.
- Rivadulla, A. (1991). *Probabilidad e inferencia científica*. Barcelona: Anthropos.
- Roach, D., Gibson, D. y Weber, K. (2004). Why is $\sqrt{25}$ not ± 5 . *Mathematics Teacher*, 97(1), 12-13.
- Rothman K. J. (1995). Causes 1976. *American Journal Epidemiology*; 141, 90-5.
- Ruiz, B. (2006). *Un acercamiento cognitivo y epistemológico a la didáctica del concepto de variable aleatoria*. Centro de investigación en ciencia aplicada y tecnología avanzada del IPN, México.

- Ruiz-Maya, L. y Martín, F. J. (2002). *Fundamentos de inferencia estadística*. Madrid: AC.
- Ruiz-Maya, L., Martín, F. J., Montero, J. M. y Uriz, P. (1995). *Análisis estadístico de encuestas: datos cualitativos*. Madrid: AC.
- Russell, B. (1973). Acerca de la noción de causa. En B. Russell. *Obras completas*. (p. 1015-1029). Madrid: Aguilar.
- San Martín, R. y Pardo, A. (1989). *Psicoestadística Contrastes paramétricos y no paramétricos*, Pirámide.
- Sánchez, J. J. (1995). *Manual de análisis de datos*, Alianza Universidad Textos.
- Sánchez, F. T. (1996). *Análisis de la exposición teórica y de los ejercicios de correlación y regresión en los textos de bachillerato*. Memoria de Tercer Ciclo. Universidad de Granada.
- Sánchez, F. T. (1999). *Significado de la correlación y regresión para los estudiantes universitarios*. Tesis Doctoral. Universidad de Granada.
- Santisteban, C. (1990). *Psicometría aplicada*. Madrid: Norma.
- Scheffé, H. (1953). A method for judging all contrast in analysis of variance. *Biometrika*, 40(1-2), 87-110.
- Schild, M. (2000). Statistical literacy: difficulties in describing and comparing rates and percentages. Trabajo presentado en el *Joint Statistical Meeting*. Atlanta, GA: American Statistical Association.
- Schild, M. (2006). Statistical literacy survey analysis: reading graphs and tables of rates and percentages. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town: International Statistical Institute and International Association for Statistical Education. On line: <http://www.stat.auckland.ac.nz/~iase>.
- Scholz, R. (1987). *Decision making under uncertainty*. Amsterdam. North Holland.
- Schuyten, G. (1991). Statistical thinking in psychology and education. En D. Vere-Jones (Ed.). *Proceeding of the Third International Conference on Teaching Statistics* (pp. 486-490). Voorburg, The Netherlands: International Statistical Institute.
- Sedlemeier, P. (1999). *Improving statistical reasoning. Theoretical models and practical implications*. Mahwah, NJ: Erlbaum.
- Shaklee, H. (1983). Human covariation judgment: accuracy and strategy. *Learning and Motivation*, 14, 433-448.

- 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.
- Shaklee, H. y Tucker, D. (1980). A rule analysis judgments of covariation between events. *Memory and Cognition*, 8, 459-467.
- Shaughnessy, J. M. (1992). Research in probability and statistic: Reflections and directions. En D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan Publishing Company.
- Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Greenwich, CT: Information Age Publishing y NCTM.
- Shaughnessy, J. M., Garfield, J., y Greer, B. (1996). Data handling. En A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, y C. Laborde (Eds.), *International handbook of mathematics education* (Vol. 1, pp. 205-237). Dordrecht: Kluwer.
- Siegel, S. (1983). *Estadística no paramétrica aplicada a las ciencias de la conducta*. Editorial Trillas. México.
- Sierpiska, A. (1994). *Understanding in mathematics*. London: Falmer Press.
- Skipper, J. K., Guenter, A. L. y Nass, G. (1970). The sacredness of .05: A note concerning the uses of statistical levels of significance in social sciences. In D. E. Morrison y R. E. Henkel, (Eds.), *The significance tests controversy: A reader* (pp. 155-160). Chicago: Aldine.
- Smedlund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, 4, 165-174.
- Solanas, A., Salafranca, L., Fauquet, J. y Núñez, M. I. (2005). *Estadística descriptiva en ciencias del comportamiento*. Madrid: Thomson.
- Tauber, L. (2001). *La construcción del significado de la distribución normal en un curso de análisis de datos*. Tesis doctoral. Universidad de Sevilla.
- Thorndike, R. L. (1989). *Psicometría aplicada*. México: Limusa.
- Tukey, J. (1953). *The problem of multiple comparisons*. Princeton University.
- Tversky, A. y Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185, 1124-1131.
- Vallecillos, A. (1994). *Estudio teórico-experimental de errores de errores y concepciones sobre al contraste estadístico de hipótesis en estudiantes universitarios*. Tesis Doctoral. Universidad de Granada

- Vallecillos, A. (1996). *Inferencia estadística y enseñanza: un análisis didáctico del contraste de hipótesis estadísticas*. Granada: Comares.
- Vallecillos, A. y Batanero, C. (1997). Conceptos activados en el contraste de hipótesis estadísticas y su comprensión por estudiantes universitarios. *Recherches en Didactique des Mathématiques*, 17, 29-48.
- Vallecillos, A. (1998). Research and teaching of statistical inference. En L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee y W. Wong (Ed.), *Proceeding of the Fifth International Conference on the Teaching of Mathematics* (pp. 296-298). Singapur: IASE.
- Vera, O. D., Díaz, C. y Batanero, C. (2011). Dificultades en la formación de hipótesis estadísticas por estudiantes de Psicología. *UNIÓN*, 27, 41-61.
- Vergnaud, G. (1990). La théorie des champs conceptuels, *Recherches en Didactique des Mathématiques*, 10(2-3), 133-170.
- Weber, R. P. (1985). *Basic content analysis*. Londres: Sage.
- Walker, H. M. (1940). Degrees of freedom. *Journal of Educational Psychology*, 31(4) 253-269.
- White, A. L. (1980). Avoiding errors in educational research. En R. J. Schumway (Ed.), *Research in mathematics education* (pp. 47-65). Reston, VA: National Council of Teachers of Mathematics.
- Wild, C., y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (con discusión). *International Statistical Review*, 67(3), 223-265.
- Williams, A. M. (1997). Students' understanding of hypothesis testing: the case of the significance concepts En: F. Biddulph y K. Karr (Eds.), *People in Mathematics Education, Proceedings of the 20th Conference of the Mathematics Education Research Group in Australasia*, (pp. 585-591). Aotearoa Australia: MERGA.
- Williams, A. M. (1998). Students' understanding of the significance level concept. En L. Pereira-Mendoza, L. Seu Kea, T. Wee Kee y W. Wong (Eds.), *Proceeding of the Fifth International Conference on the Teaching of Mathematics* (pp. 743-749). Singapore: IASE.
- Wright, J. C. y Murphy, G. L. (1984). The utility of theories in intuitive statistics: the robustness of theory-based judgments, *Journal of Experimental Psychology General*, 113(2), 301-322.
- Yates, J. F. y Curley, S. P. (1986). Contingency judgment: Primacy effects and attention decrement, *Acta Psicológica*, 62, 293-302.



ANEXO 1
CUESTIONARIO DE EVALUACIÓN DE
CONOCIMIENTOS PREVIOS





PRUEBA DE COMPRENSIÓN DE LA NOCIÓN DE ASOCIACIÓN
ESTADÍSTICA

Nota de acceso en selectividad:.....

Especialidad de Bachillerato:.....

Nombre y apellido:.....

¿Repite la asignatura? :.....

Lee atentamente los enunciados y responde las preguntas planteadas, dando detalle de tus razonamientos. Te agradecemos tu colaboración.

1. Se pretende estudiar si el sufrir insomnio tiene relación con los trastornos de estrés. En una muestra de 250 personas observadas se obtuvieron los siguientes resultados:

	Padece estrés	No padece estrés
Tener insomnio	90	60
No tener insomnio	60	40

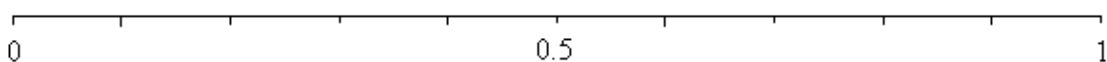
a. Utilizando estos datos razona si tener o no insomnio tiene relación con padecer estrés.

b. Indica como has usado los datos de la tabla para llegar a tu conclusión

c. Asigna una puntuación entre 0 (mínimo) y 1 (máximo) según la intensidad de esta relación, marcando una cruz en el punto de esta escala que creas adecuado:

0: Indica que no existe ninguna relación entre tener insomnio y padecer estrés

1: Indica que sabiendo si una persona padece o no insomnio podemos decir con seguridad si tiene o no estrés



2. Un psicólogo infantil estudia si existe relación entre ser hijo único (sin hermanos) y ser un niño problemático. Para ello se han observado 250 niños obteniendo los resultados siguientes

	Niño problemático	Niño no problemático
Tiene hermanos	40	100
Hijo único	100	10

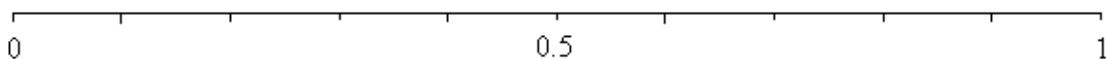
a. Utilizando estos datos razona si hay alguna relación entre ser un niño problemático y tener o no hermanos.

b. Indica cómo has utilizado los números de la tabla, para llegar a tu conclusión.

c. Asigna una puntuación entre 0 (mínimo) y 1 (máximo) según la intensidad de esta relación, marcando una cruz en el punto de esta escala que creas adecuado:

0: Indica que no existe ninguna relación entre tener hermanos y ser un niño problemático

1: Indica que sabiendo si un niño tiene o no hermanos podemos decir con seguridad si es problemático o no



3. Se quiere saber si sufrir o no de alergia tiene relación con llevar una vida sedentaria (llevar una vida sin realizar ningún tipo de ejercicio físico). Para ello obtuvimos los siguientes resultados de 300 sujetos:

	Sufre alergia	No sufre alergia
Forma de vida sedentaria	130	30
Forma de vida no sedentaria	20	120

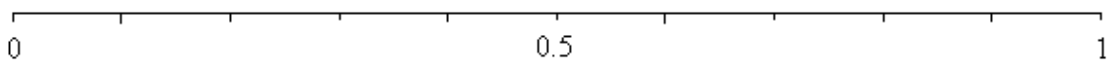
a. Utilizando estos datos razona si padecer alergia tiene o no relación con llevar una vida sedentaria en estos sujetos.

b. Indica como has usado los números de la tabla, para obtener tu conclusión.

c. Asigna una puntuación entre 0 (mínimo) y 1 (máximo) según la intensidad de esta relación, marcando una cruz en el punto de esta escala que creas adecuado:

0: Indica que no existe ninguna relación entre sufrir alergia y llevar una forma de vida sedentaria

1: Indica que sabiendo si una persona padece o no de alergia podemos decir con seguridad si lleva o no una forma de vida sedentaria



4. La siguiente tabla nos indica el número de estudiantes que aprobaron o suspendieron un examen teniendo en cuenta el tiempo que cada estudiante dedicó a prepararlo

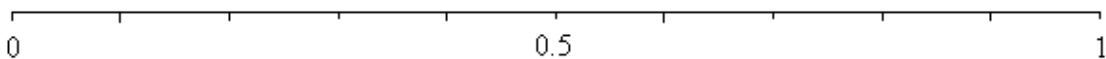
	Menos de 5 h	Entre 5 y 10 h	Más de 10 h	TOTAL
APROBADOS	5	15	51	71
SUSPENSOS	20	7	2	29
TOTAL	25	22	53	100

a. Utilizando estos datos razona si aprobar o suspender el examen tiene relación con el tiempo dedicado a prepararlo.

b. Indica como has usado los números de la tabla, para llegar a tu conclusión.

c. Asigna una puntuación entre 0 (mínimo) y 1 (máximo) según la intensidad de esta relación, marcando una cruz en el punto de esta escala que creas adecuado:

0: Indica que no existe ninguna relación entre aprobar y el número de horas dedicadas
 1: Indica que sabiendo el número de horas dedicadas podemos decir con seguridad si el estudiante aprobará o no



MUCHAS GRACIAS POR TU COLABORACIÓN



ANEXO 2
EJERCICIOS PROPUESTOS EN LAS
SESIONES PRÁCTICAS Y
EVALUACIONES



1. PRIMERA SESIÓN DE PRÁCTICAS

Actividad 1. En un colegio se pregunta a los alumnos, obteniendo los siguientes resultados:

	Chicos	Chicas	Total
Le gusta el tenis	400	200	600
No le gusta	50	50	100
Total	450	250	700

Si elegimos al azar uno de estos alumnos:

- a. ¿Cuál es la probabilidad de que le guste el tenis?
- b. ¿Cuál es la probabilidad de que sea una chica y además le guste el tenis?
- c. Sabiendo que el alumno elegido es chica, ¿Cuál es la probabilidad de que le guste el tenis?

Actividad 2. Un psicólogo en un estudio obtuvo la siguiente clasificación de los pacientes de un hospital de acuerdo con las variables “sexo” y “diagnóstico”.

Sexo	Diagnóstico	
	Esquizofr.	No esquizofr.
Varón	43	15
Mujer	32	52

- a. ¿Hay asociación entre las variables? ¿Es directa o inversa?
- b. ¿Por qué?

Actividad 3. Para comprobar si existía relación entre el tabaco y la hipertensión, fueron elegidas aleatoriamente 100 personas, de las que 55 resultaron ser hipertensas.

Agrupadas dichas personas en tres clases, según que no fumaran, fumaran moderadamente o fumaran mucho, se encontró que entre los hipertensos 10 no fumaban y 25 lo hacían moderadamente. En cambio, entre los no hipertensos 18 fumaban moderadamente y 7 fumaban mucho.

- a. Organiza los datos para formar una tabla de doble entrada o de contingencia.
- b. Calcula las distribuciones condicionales por filas.
- c. Dibuja el diagrama de barras adosado y diagrama de barras apilado.

Actividad 4. Un psicólogo en un estudio obtuvo la siguiente clasificación de una muestra de 100 sujetos de acuerdo con las variables, “concepción de la inteligencia” y “tipo de autoinstrucciones”

Concepción Inteligencia	Autoinstrucción		
	Respuestas instruccionales	Respuestas atribucionales	Otras respuestas
Destreza	24	11	9
Rasgo	72	33	27

Comprueba si se cumplen las propiedades de independencia

2. SEGUNDA SESIÓN DE PRÁCTICAS (ANALIZADOS EN EL CAPÍTULO 6)

Problema A. Para un experimento sobre memoria a largo plazo un psicólogo selecciona 3 palabras y decide evaluar el componente emocional de las mismas con una muestra aleatoria. Presenta, independientemente, cada palabra a 100 sujetos y registra si la palabra es percibida como emocionalmente positiva (agradable, placentera, etc.), negativa (desagradable, displacentera, etc.) o neutra. A la vista de los resultados obtenidos, ¿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.

Problema B. Se desea estudiar hasta qué punto existe relación entre el tiempo de residencia de inmigrantes en nuestro país y su percepción de integración. Se dispone de una muestra de 207 inmigrantes a los que se les evaluó en ambas variables obteniéndose la siguiente tabla de frecuencias observadas.

Tiempo de Residencia	Grado de integración		Total
	Bajo	Alto	
Más tiempo	36	81	117
Menos tiempo	81	9	90
Total	117	90	207

Calcule e interprete las medidas de asociación para tablas 2x2

Problema C. Se está realizando un estudio sobre la población de estudiantes que hicieron la prueba de selectividad en el curso 1978-1979. Dos de las variables registradas fueron “convocatoria” (junio-septiembre) y “tipo de estudios” por el que se inclinaba el estudiante (grado medio, facultades, escuelas técnicas). Los resultados encontrados en la Universidad Complutense fueron los expuestos en la tabla siguiente:

	Escuelas Técnicas	Facultades	Grado Medio
	Septiembre	200	500
Junio	1500	1300	700

Calcule e interprete las medidas de asociación para tablas rxc

3. EJERCICIOS PROPUESTOS EN LAS EVALUACIONES (ANALIZADOS EN EL CAPÍTULO 6)

Problema D. Un grupo de 200 personas aquejadas de ansiedad fue dividido aleatoriamente en dos subgrupos. Al primer grupo se ofreció unas píldoras realmente efectivas para que tomaran tres al día, y al otro se ofreció un placebo (medicamento sin efecto). Al cabo de un mes fueron interrogados sobre la eficacia de las pastillas tomadas, con el siguiente resultado:

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

- a. Preguntado un paciente, dice que su ansiedad casi ha desaparecido; está encantado con el tratamiento. ¿Cuál es la probabilidad de que este paciente haya tomado el placebo?
- b. ¿Hay asociación entre el tipo de tratamiento (medicamento o placebo) y el efecto producido (la ansiedad disminuye o no)? ¿O son las variables independientes?
- c. Indica cómo has llegado a esta conclusión (puedes usar el método que prefieras)

Problema E. En un estudio se analiza la relación entre la dieta pobre en calcio y el desarrollo de osteoporosis. Se obtienen los siguientes resultados

	Desarrolla osteoporosis	
Dieta	Si	No
Pobre en calcio	58	62
Normal	22	258

- a. ¿Hay asociación entre el tipo de dieta (rica o pobre en calcio) y el desarrollo de osteoporosis?
- b. ¿Es la relación directa o inversa?
- c. ¿Es fuerte o moderada?
- d. Utiliza algún coeficiente de asociación que recuerdes para justificar tus respuestas
- e. Si no recuerdas los coeficientes, da la justificación usando cualquier propiedad que hayas aprendido en el temas de tablas de contingencia



ANEXO 3
CUESTIONARIOS DE EVALUACIÓN
DEL APRENDIZAJE



1. CUESTIONARIOS DE EVALUACIÓN DE LOS TEMAS 1 Y 2

- A1. La principal diferencia entre el diagrama de barras apilado y el adosado es:
- El diagrama de barras apilado sólo se puede construir con frecuencias absolutas, mientras que el adosado se puede construir con frecuencias absolutas o porcentajes
 - En el diagrama de barras apilado las frecuencias condicionales de cada valor de X para un mismo valor de Y se representan en una misma barra y en el adosado se representan una al lado de la otra.
 - En el diagrama de barras apilado las frecuencias marginales de cada valor de X se representan en una misma barra y en el adosado se representan una al lado de la otra
- A2. En una tabla de contingencia, la suma de las frecuencias relativas marginales por filas es:
- Igual a la suma de frecuencias relativas marginales por columnas
 - Igual al total de la muestra
 - Igual a la suma de las frecuencias absolutas marginales dividida por el total de la muestra
- A3. Las frecuencias dobles absolutas y relativas están relacionadas entre sí mediante:
- Las frecuencias condicionales.
 - Las frecuencias marginales.
 - El tamaño de la muestra.
- A4. Para calcular las frecuencias relativas condicionadas es necesario:
- Sólo las frecuencias absolutas dobles de la fila o columna por la que se condiciona.
 - Las frecuencias relativas de la fila o columna por la que se condiciona y el total de la muestra.
 - Las frecuencias absolutas o relativas de la fila o columna por la que se condiciona.
- A5. Las tablas rxc se pueden representar gráficamente en los gráficos:
- Únicamente en el diagrama de barras apilado y el adosado.
 - Únicamente en el gráfico tridimensional y el diagrama de barras adosado.
 - En los mismos gráficos que una tabla 2x2.
- A6. El objetivo de construir una tabla de contingencia es:
- Resumir los datos de dos variables cuantitativas
 - Resumir los datos de dos variables cualitativas
 - Resumir los datos de dos variables cualitativas, o bien numéricas con pocos valores diferentes

B1. Para que dos variables de una tabla de contingencia sean independientes han de ser iguales:

- Las frecuencias relativas condicionales por columnas.
- Las frecuencias relativas condicionales por filas.
- Las frecuencias relativas condicionales y frecuencias relativas marginales.

B2. Las frecuencias esperadas se calculan mediante:

- Las frecuencias absolutas dobles y el total de la muestra.
- Las frecuencias relativas dobles y el total de la muestra.
- Las frecuencias absolutas marginales y el total de la muestra.
- Las frecuencias relativas marginales

B3. En caso de que haya asociación entre variables, las frecuencias relativas dobles:

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

B4. En las siguientes tablas 2x2 indicamos el tipo de asociación que informan las diferentes celdas ¿Cuales de las siguientes tablas es correcta?

	a.		b			c		
	B	No B	A	B	No B	A	B	No B
A	Dep. directa	Dep. directa	A	Dep. directa	Dep. inversa	A	Dep. directa	Dep. inversa
No A	Dep. directa	Dep. inversa	No A	Dep. inversa	Dep. directa	No A	Dep. inversa	Dep. directa

B5. Indica cuál de las siguientes frases es cierta:

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

B6. La diferencia entre la dependencia funcional y la dependencia aleatoria consiste en:

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

2. CUESTIONARIOS DE EVALUACIÓN DE LOS TEMAS 3 Y 4

- C1. Para realizar un test Chi- cuadrado en una tabla de contingencia con m filas y n columnas, los grados de libertad del estadístico chi-cuadrado son:
- a. $m \times n$. b. $m + n$. c. $(m-1) \times (n-1)$. d. $(m-1) + (n-1)$.
- C2. Los grados de libertad sirven para:
- a. Contar el número de filas y columnas.
b. Calcular el parámetro de la distribución del estadístico.
c. Encontrar en las tablas de la distribución la probabilidad de un valor de Chi-cuadrado, y saber si el valor de Chi cuadrado es estadísticamente significativo.
- C3. Un investigador usa un test Chi- cuadrado para determinar si hay asociación entre dos variables en una tabla de contingencia. Calcula el valor de Chi cuadrado y obtiene el valor p correspondiente. ¿Cuál de los siguientes valores p indicaría de forma más concluyente que hay asociación entre las variables.
- a. .002 b. .01 c. .05 d. .10
- C4. Un investigador quiere estudiar si existe relación entre dos variables. Calcula el estadístico Chi- cuadrado y obtiene un valor p estadísticamente significativo ($p < 0,05$). La decisión que debe tomar es:
- a. Rechazar la hipótesis de independencia entre las dos variables
b. Rechazar la hipótesis de asociación entre las dos variables
c. Rechazar la hipótesis de asociación directa entre las dos variables
d. Con estos datos, no podemos saber la decisión que se debe tomar
- C5. Un investigador quiere estudiar la relación entre dos variables estadísticas. Usando un programa de ordenador obtiene los siguientes valores en un contraste Chi cuadrado

	Valor	gl	Significación (valor p)
Chi-cuadrado de Pearson	157,331	4	,005
N de casos válidos	398		

0 casillas (,0%) tienen una frecuencia esperada inferior a 5.
La frecuencia mínima esperada es 14,77.

- a. El contraste es estadísticamente significativo, pero el investigador no puede aplicar el contraste Chi- cuadrado porque ninguna celda tiene frecuencia esperada menor que 5
- b. El contraste es estadísticamente significativo, y el investigador puede aplicar el contraste Chi- cuadrado porque se cumplen las condiciones
- c. El contraste es estadísticamente significativo, pero el investigador no puede aplicar el contraste Chi- cuadrado porque la frecuencia mínima esperada es menor de 20.
- C6. El investigador anterior:
- a. Debe concluir que no hay asociación entre las dos variables en su estudio
- b. Debe concluir que hay asociación entre las dos variables en su estudio, y la probabilidad de que las variables en realidad sean independientes es menor que 0,005
- c. Debe concluir que hay asociación entre las dos variables en su estudio; la probabilidad de obtener el valor $\chi^2 = 157,331$ si las variables son en realidad independientes es menor que 0,005

- D1. Si el coeficiente Phi tiene un valor 0,3, esto indica
- Relación inversa y moderada entre las variables
 - Relación directa y moderada entre las variables
 - Relación directa y alta entre las variables
 - Relación inversa y alta entre las variables
- D2. Las medidas de la asociación para tablas 2x2 son:
- Coeficiente Phi de Pearson, riesgo relativo y coeficiente de contingencia de Pearson.
 - Coeficiente Phi de Pearson, riesgo relativo y razón de productos cruzados.
 - Riesgo relativo, razón de productos cruzados y V de Cramer.
- D3. El riesgo relativo por filas y columnas:
- Difieren y el valor por filas siempre es mayor o igual que el valor por columnas.
 - Son diferentes su forma de cálculo, pero coinciden.
 - Difieren y no hay una regla para saber cual es mayor.
- D4. La razón de productos cruzados varía de:
- $0 < RC < 1$.
 - $-1 < RC < 1$.
 - Puede ser mayor o menor que 1
- D5. Un investigador quiere relacionar el número de cilindros de una serie de marca de automóviles con el sitio donde se han fabricado. Obtiene las siguientes medidas de asociación en sus datos:

	Valor
Lambda Número de cilindros dependiente	,183
Lambda País de origen dependiente	,000
V de Cramer	,445
C Coeficiente de contingencia	,532

- Los valores de Lambda indican que se obtiene mejor predicción al predecir el número de cilindros sabiendo el país de origen que al predecir el país de origen sabiendo el número de cilindros.
 - Los valores de Lambda indican que se obtiene mejor predicción al predecir el país de origen sabiendo el número de cilindros que al predecir el número de cilindros sabiendo el país de origen.
 - Se obtiene igual predicción en ambos casos
- D6. En el ejemplo anterior, los valores de los coeficientes de asociación V y C indican:
- Una asociación positiva débil
 - Una asociación positiva moderada
 - Una asociación negativa moderada
 - Tendríamos que realizar el contraste Chi-cuadrado para conocer la intensidad y el signo

3. PREGUNTAS INCLUIDAS EN LA EVALUACIÓN FINAL DE LA ASIGNATURA

E1. Se quiere estudiar la posible asociación entre que una madre gestante fume durante su embarazo y que el niño sea bajo de peso al nacer. En un estudio se obtienen los siguientes datos:

	Recién nacido de bajo peso	
	Sí	No
Gestante Fumadora	43	207
No fumadora	105	1645

¿Qué conclusiones puede obtener?

- Hay una relación directa entre las variables; se puede observar al comparar las celdas que están en las diagonales sin necesidad de calcular el estadístico chi-cuadrado.
- No hay relación pues la mayoría de las madres tiene un niño de peso normal y no fuma durante el embarazo.
- Hay relación inversa pues la mayoría de las madres fumadoras tiene un niño de peso normal y son muy pocas las madres fumadoras con niños bajos de peso.
- No cumple las propiedades de aplicación del estadístico Chi-cuadrado, por lo tanto no podemos dar ninguna conclusión referida a la dependencia.

E2. Un psicólogo pone a punto una terapia para disminuir la ansiedad. Recoge datos de 200 pacientes a 100 de los cuáles (elegidos al azar) aplica la terapia y al resto otra alternativa y pasados tres meses de tratamiento analiza la posible asociación entre terapia seguida (nueva o antigua) y disminución del nivel de ansiedad (disminuye o no disminuye). Aplica un contraste Chi-cuadrado de independencia y obtiene un valor $\chi^2=49,1$, con 1 g.l. ¿Qué conclusión debiera obtener de estos resultados?

- El resultado del contraste no es estadísticamente significativo, pues el valor en la tabla así lo confirma al compararlo con el obtenido por el investigador. Debiera concluir que las variables no están relacionadas entre sí.
- El resultado del contraste es estadísticamente significativo, pues $p < 0,05$. Debiera concluir que las variables están relacionadas entre sí. La relación es muy pequeña, porque el valor p es muy bajo.
- El resultado del contraste es estadísticamente significativo, pues $p < 0,05$. Debiera concluir que las variables están relacionadas entre sí. Para saber si la relación es fuerte o débil tenemos que calcular alguna medida de asociación.
- No podemos concluir si hay relación entre las variables sólo con el valor Chi-cuadrado. Tenemos que calcular los coeficientes de asociación.

E3. Un investigador quiere relacionar el número de cilindros de una serie de marca de automóviles con el sitio donde se han fabricado (Europa, Estados Unidos o Japón). Obtiene las siguientes medidas de asociación en sus datos

V de Crámer= 0,445; Coeficiente de contingencia de Pearson= 0,532, Lamda de Goodman y Kruskal (Variable dependiente= Número de cilindros)= 0,183

¿Qué conclusión debe obtener el investigador?

- a. Hay una relación muy alta entre las variables y se puede predecir muy fácilmente el número de cilindros del coche, si se sabe donde se ha fabricado
- b. Hay una relación moderada-alta entre las variables y se puede predecir muy fácilmente el número de cilindros del coche, si se sabe donde se ha fabricado
- c. Hay una relación moderada-alta entre las variables, pero no es fácil predecir el número de cilindros del coche, si se sabe donde se ha fabricado
- d. Hay una relación muy alta entre las variables, pero no es fácil predecir el número de cilindros del coche, si se sabe donde se ha fabricado

E4. Dos variables en una tabla de contingencia son independientes entre sí. Indica cuál de las siguientes propiedades NO SE CUMPLE

- a. Las frecuencias observadas de todas las celdas son iguales a las frecuencias teóricas esperadas.
- b. En todas las celdas son iguales al producto del total por fila y columna que le corresponda, es decir $h_{i,j} = h_i \cdot h_j$
- c. La suma de las frecuencias de las celdas en la diagonal principal será mayor que la suma de frecuencias en las celdas que no están en esta diagonal.
- d. Las frecuencias relativas condicionales por filas son iguales a las frecuencias relativas marginales por filas.



ANEXO 4
RESUMEN DE LA OBSERVACIÓN DE
LA ENSEÑANZA



Sesión 1. Tema 1. Tablas de contingencia. Lectura e interpretación

Se entrega a los alumnos una copia del cuadernillo de apuntes y se les explica la finalidad del curso, la forma en que se va a trabajar, dentro del mismo, así como los materiales que se han preparado. Se les pasa la lista para controlar la asistencia y se les da información sobre la página web, donde se han colocado los materiales del curso y el correo electrónico del profesor que pueden utilizar para hacer preguntas sobre los aspectos que no hayan comprendido. Se explican los objetivos del primer tema y el interés del mismo para resumir los datos en forma de tabla de contingencia, identificar las frecuencias que corresponde a cada par de valores de las variables, calcular las frecuencias relativas dobles, marginales y condicionales absolutas y relativas e interpretarlas.

Se inicia el tema con la introducción, recordando el significado del concepto de variable estadística, la diferencia entre valor y variable y los tipos de variable (cuantitativo y cualitativo). Se introduce la tabla de contingencia a partir del Ejemplo 1.1 (tabla 2x2) sobre problemas neuronales y tratamiento con dos fármacos diferentes que se quieren comparar. Se explica el significado y notación de las variables, y una vez comprendido el problema, se introduce el diagrama de barras adosado, utilizando para ello el mismo ejemplo.

- El profesor pide leer una frecuencia absoluta doble en la tabla 1.2. Los alumnos responden correctamente identificando un valor de la tabla y a qué categorías de las variables corresponde.

Se sigue introduciendo las frecuencias dobles absolutas y relativas, justificando el interés del uso de frecuencias relativas. Se explica la notación y cálculo de las frecuencias relativas, así como la relación entre frecuencias absolutas y relativas y la propiedad de la suma de cada una de ellas mediante el ejemplo Tabla 1.3.

- Se pregunta por el significado de una frecuencia relativa en la tabla 1.3. Nadie contesta, pues no identifican el porcentaje a partir de la frecuencia relativa. El profesor lo explica y la forma de pasar de frecuencia relativa a porcentaje.
- Preguntan por qué se pone h en vez de f para identificar las frecuencias relativas. El profesor lo explica.

Se continúa la explicación de las frecuencias marginales, su notación y cálculo. Se hace ver que la distribución marginal corresponde a una sola variable, mientras que la doble corresponde a dos variables conjuntas. Se hace ver que a partir de la tabla doble se pueden generar dos distribuciones marginales (por filas y columnas).

Se sigue con la explicación del significado y cálculo de las frecuencias condicionales relativas. Se hace ver que existen varias distribuciones condicionales por filas y por columnas. Se explica la relación entre frecuencias condicionales, dobles y marginales, así como la suma de frecuencias condicionales por filas y columnas. Todas las explicaciones se apoyan en el ejemplo para mejorar la comprensión.

- Un alumno pregunta por el valor que hay que dividir las frecuencias absolutas para obtener las frecuencias condicionales. El profesor lo explica.

Finalizadas todas estas explicaciones, se indica que el procedimiento de cálculo de los distintos tipos de frecuencias es similar para las tablas rxc y se generaliza la notación adecuada.

Se introducen los diagramas de barras adosados, diagramas de barras apilados y gráfico tridimensional a partir del mismo ejemplo, comparando la información que proporcionan unos

y otros gráficos. Después de finalizar el tema se hace un resumen de los principales conceptos. Los alumnos no hacen preguntas durante este tiempo.

A continuación el profesor propone completar el Ejercicio 1.1 sobre despersonalización y relaciones personales.

- Se comienza interpretando las frecuencias dobles de la tabla dada en el enunciado del ejercicio.
- Se pregunta a un alumno A1 el significado de una celda. Responde, pero no define correctamente las variables, pues indica solo los valores “muchas” “bajas” en lugar de decir “es el número de personas con muchas relaciones personales y baja despersonalización”.
- Los alumnos A2 y el A3 por el contrario, preguntados por otros valores de las celdas hacen una lectura correcta.

Se comienza el apartado 1 “Interpreta las frecuencias absolutas marginales”.

- El alumno A1 define correctamente los datos.
- La alumna A2 no sabe responder, “está perdida”. El profesor la ayuda a situarse.
- Una alumna pregunta que son las frecuencias marginales y la distribución marginal y se le responde.

Se sigue con el apartado 2 “Calcula e interpreta las frecuencias relativas dobles”.

- Se pregunta al alumno A1, quien las calcula e interpreta correctamente las cuatro frecuencias.

Se continúa con el apartado 3 “Frecuencias marginales relativas”.

- El alumno A1 las calcula correctamente, algunos alumnos aprueban este cálculo, pues están de acuerdo.
- El alumno A2 calcula correctamente, la mayoría está de acuerdo.
- El alumno A3 calcula correctamente.
- Una alumna pregunta que son las frecuencias marginales y la distribución marginal y se le responde.

El profesor inicia el apartado 4: Interpretar las frecuencias relativas condicionales por filas.

- Los alumnos A1, A2, A3 y A4, interpretan correctamente cada uno una celda diferente.

Se sigue con el apartado 5: Interpretar las frecuencias relativas condicionales por columnas.

- El alumno A1 calcula correctamente.

No hubo tiempo para hacer la pregunta sobre cálculo de probabilidades (segundo apartado del ejercicio) y se pide a los estudiantes que lean los apuntes donde viene explicada la solución. El segundo ejercicio se explica sólo los pasos en forma general, sin entrar en detalle

de cómo se resuelve paso a paso. Se pide a los estudiantes que lean los apuntes donde viene explicada la solución con detalle.

Sesión 2. Tema 2. Asociación estadística, dependencia funcional e independencia

Se recuerda a grandes rasgos lo visto el día anterior, entablando una relación con el siguiente tema. Se les pasa la lista para controlar la asistencia y se les recuerda la existencia de la página web, con los materiales y el correo electrónico del profesor para hacer preguntas. Se introducen los objetivos del tema, que son diferenciar la asociación estadística, dependencia funcional e independencia; reconocer el tipo de relación entre dos variables de una tabla comparando las frecuencias condicionales, Calcular las frecuencias esperadas en caso de independencia y analizar posibles explicaciones de una asociación estadística

Se explican los distintos tipos de dependencias (funcional y aleatoria), y el concepto de asociación (independencia, asociación parcial y asociación perfecta). Se utiliza el ejemplo del apartado 2.2 sobre velocidad, espacio y tiempo para explicar la dependencia funcional. Se habla de los diagramas de dispersión y la dispersión de los puntos para comparar la dependencia aleatoria con la funcional.

Se empieza la explicación referida a la independencia en tablas de contingencia, basándose en el ejemplo 2.1 sobre trastornos digestivos. Se discute la Figura 2.2 donde se presentan diferentes tipos de asociación (independencia, parcial y perfecta) en tablas de contingencia, diferenciando entre estos tipos de asociación e indicando que lo usual es el caso de asociación parcial. También se indica que en los temas que siguen se estudiará la forma de medir la intensidad de la asociación y ver si es estadísticamente significativa.

Se comienza a analizar la idea de independencia con el ejemplo 2.2.1 cual vieron los alumnos en el cuestionario anterior a la enseñanza (sobre relación entre Insomnio y estrés). A pesar de que las teorías previas de los alumnos les llevan a pensar que debe haber relación entre las variables, los datos de la tabla se han preparado para que haya independencia perfecta.

Siguiendo el ejemplo, se calculan las frecuencias relativas condicionales por columnas, comprobando que las dos distribuciones resultantes son iguales entre sí e igual a la distribución de frecuencias relativas marginal por columna. Se hace lo mismo con las frecuencias relativas condicionales por filas, comprobando que también son iguales entre si y con la marginal correspondiente.

El profesor indica que cuando ocurre esto, las variables son independientes e introduce en forma general y utilizando la notación dicha propiedad, es decir verifica para todo par de valores i, j . $h(x_i/y_j) = h_i$. Y $h(y_j/x_i) = h_j$

- Una alumna pregunta por el significado de las h_{ij} (la alumna faltó el día anterior). El profesor lo comenta que se refiere a frecuencias relativas, y comenta también el significado de las frecuencias dobles f_{ij} .
- Se termina de explicar la independencia en tablas de contingencia y la propiedad que La frecuencia relativa doble es igual al producto de las frecuencias relativas marginales de su fila y su columna: $h_{ij} = h_i \cdot h_j$, para todo i, j

Se hace un resumen de todas las propiedades de la frecuencia en caso de independencia, las cuales las comprobamos en el mismo ejemplo (Insomnio-estrés).

- Una alumna dice: “*Yo es que no me entero*” (sobre propiedades de la independencia), el profesor retrocede y lo explica nuevamente más despacio hasta que la alumna parece entenderlo.

Como final de la parte teórica se introduce el punto 2.7 sobre el caso particular de la tabla 2x2 y sobre las celdas que informan de dependencia directa e inversa. Se utiliza el ejemplo 2.4 sobre alergia y vida sedentaria. Se finaliza con el punto 2.6 sobre posibles explicaciones de la asociación. Se recuerda que la asociación puede ser debida a la existencia de una relación causal entre las variables, pero que existen otras explicaciones y que correlación no implica causalidad. Se les pone el ejemplo sobre libros y coeficiente intelectual.

- Una alumna pregunta sobre la diferencia ente dependencia directa e inversa en una tabla de contingencia 2x2, y su interpretación, el profesor lo explica poniendo un ejemplo hasta que la alumna parece entenderlo.

Después de finalizar el tema se hace un resumen Los alumnos no hacen preguntan durante este tiempo. A continuación el profesor propone resolver el Ejemplo 2.4 (vida sedentaria-sufrir alergia) en voz alta para comprobar si hay independencia

- Se pregunta a un alumno, el cual no entiende lo que hay que hacer. Se le aclara el enunciado y responde correctamente.
- Se pregunta a otro alumno, responde correctamente.

El profesor propone resolver el ejercicio 2.1 (sobre decisión en dos colegios), para el cual deja un poco de tiempo para que se haga. Se trata de ver si hay o no dependencia entre las variables.

- Se pregunta a un alumno, el cual no sabe hacerlo, por no haber venido el día anterior. Se le explica como se puede hacer, el profesor lo anima a que lo intente y le pregunte dudas.
- Se pregunta un alumno si tiene las frecuencias relativas condicionales por filas, el alumno las ha calculado correctamente.
- El profesor pregunta a otra alumna las frecuencias condicionales por columnas, y la alumna las da correctamente.
- Entre el profesor y los alumnos terminan el ejercicio.

A continuación el profesor propone resolver el ejercicio 2.2 (caries dental en niños y padres), para el cual deja un poco de tiempo para que se haga.

- Un alumno pregunta ¿Qué son las variables ordinales?, el profesor indica que no influye en el problema, pero luego le contestará con más detalle.
- Los alumnos encuentran un fallo en los cálculos del ejercicio en los apuntes que tienen este ejercicio hecho, lo cual retrasa un poco la explicación, pero indica que los alumnos están comprendiendo el tema, pues han sido capaces de encontrar el error.

- Para solventar la situación el profesor corrige el resultado en la presentación y comenta el resultado correcto un poco.

El profesor sugiere resolver el ejercicio 2.3 (infarto de miocardio), para el cual deja un poco de tiempo para que se haga.

- Una alumna pregunta si se puede resolver por filas, y el profesor dice que si.
- Una alumna pregunta otra vez el significado de las f y h, el profesor explica la diferencia.
- El profesor al ver que queda poco tiempo y que hay alumnos que al no venir el día anterior van más lentos en la resolución, decide poner las diapositivas con la solución mientras lo comenta.

Este día se pudo finalizar todo el tema. El profesor pide a los estudiantes que lean con detalle los apuntes en su casa.

Sesión 3. Tema 3. El estadístico Chi-cuadrado y contrastes asociados

Se recuerda a grandes rasgos lo visto el día anterior, para poder establecer una relación con el nuevo tema, donde se pretende formalizar un poco más lo aprendido hasta el momento.

Se les pasa la lista para controlar la asistencia y se les recuerda la existencia de la página web, con los materiales y el correo electrónico del profesor para hacer preguntas. Se introducen los objetivos del tema, que son aprender a obtener una medida de la diferencia entre frecuencias observadas y esperadas en caso de independencia en una tabla de contingencia; calcular e interpretar el estadístico Chi-cuadrado y sus grados de libertad; recordar los pasos y conceptos básicos en un contraste de hipótesis; comprender los pasos para llevar a cabo el contraste de independencia y aplicarlo en algunos ejemplos, decidiendo si el valor Chi-cuadrado es estadísticamente significativo y tomando una decisión sobre el rechazo o de la hipótesis nula. Igualmente para el contraste de homogeneidad para comprobar si varias subpoblaciones vienen de la misma población; Interpretar el nivel de significación y comprender los supuestos de aplicación del contraste Chi-cuadrado

Los alumnos ya habían estudiado previamente el contraste de hipótesis, en asignaturas previas, realizando contrastes de diferencias de medias y proporciones, por lo que ya conocen los fundamentos y conceptos básicos en el contraste de hipótesis. Así que el profesor comienza indicando que en este tema se van a estudiar dos tipos especiales de contraste a partir de los datos de contingencia, el *contraste de homogeneidad* y el *contraste de independencia* que se basarán en el estadístico Chi-cuadrado.

Para introducirlo se utiliza el ejemplo del apartado 3.1 sobre supervivencia en el Titanic, adelantando que se va a necesitar las frecuencias esperadas explicadas en el tema anterior, para lo cual el profesor recuerda como se calculan $e_{ij} = \frac{f_i \cdot f_j}{n}$. A continuación se introduce el estadístico, indicando que proporciona una medida de diferencia entre las frecuencias esperadas y las frecuencias observadas. Se indica que el estadístico se calcula en la forma siguiente:

$$\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}, \text{ y lo calcula en el mismo ejemplo. Se comentan las propiedades del}$$

estadístico

- Si todas las frecuencias observadas son iguales a la correspondiente frecuencia esperada, toma un valor nulo y esto ocurre sólo cuando las dos variables de la tabla son independientes.
- Cuanto mayor sea la diferencia entre las frecuencias observadas y esperadas en la tabla, el valor de Chi cuadrado será mayor. Es decir, a mayor intensidad de la asociación entre las variables, Chi-cuadrado será mayor.

Se explica el concepto de *grados de libertad* y la forma de calcularlo. A continuación se discute la lectura de la tabla de Chi-cuadrado, para obtener a partir de los grados de libertad y del valor empírico del estadístico el p-valor.

Se presenta acá un pequeño problema, ya que en los apuntes la tabla proporcionada corresponde a un calor alfa a la derecha, es decir al percentil del $(1-\alpha)\%$, pero los alumnos en las tablas que están acostumbrados a manejar de este estadístico, el alfa es a la izquierda, es decir, las tablas corresponden al percentil de $\alpha\%$.

- Un alumno pregunta sobre este punto, mostrando su extrañeza sobre el formato de la tabla. El profesor le aclara la existencia de diversos tipos de tabla y le indica que para la prueba de evaluación se les proporcionará las tablas a las que están habituados.

El profesor continúa con el mismo ejemplo, para mostrar como se calculan los grados de libertad, indica que el alfa a considerar en caso de que el problema no indique otra cosa será de 0,05 y explica el valor que se obtendría en este caso en nuestras tablas. Finalmente dedica un tiempo a la interpretación del valor crítico presentado en la tabla, indicando que, para los grados de libertad dados hay una probabilidad igual a 0,05 de obtener un de Chi cuadrado igual o mayor al valor crítico, en caso de que las variables sean independientes.

En el ejemplo, para $\alpha=0,05$ se obtiene un valor crítico de 5,99. Por otro lado el valor de Chi-cuadrado obtenido en el ejemplo 1.1 es igual a $\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 174,4$ mucho mayor que el crítico

Se indica que el valor obtenido 174,4 es muy poco probable para 2 grados de libertad en caso de independencia, pues la probabilidad de obtener un valor mayor que 10,6 es sólo 0,005. Deducimos que el salvamento de los viajeros en el Titanic no fue independiente de su clase social.

Terminado el ejemplo, el profesor repasa lo que se ha aprendido sobre el contraste de independencia, mostrando un resumen de los pasos que hay que seguir:

1. Fijar las hipótesis que se quieren contrastar, que son las siguientes:
 H_0 : Las variables en filas y columnas de la tabla son independientes
 H_1 : Hay asociación entre las filas y columnas de la tabla
2. Fijar el nivel de significación; si no se indica nada se toma $\alpha=0,05$. Se recuerda el significado del nivel de significación o probabilidad máxima que fijamos para el error tipo I (rechazar la hipótesis de independencia cuando sea falsa)
3. Calcular el estadístico Chi cuadrado:
 - Si el valor calculado χ_{exp}^2 tiene una probabilidad menor que α (nivel de significación) rechazamos la hipótesis nula H_0 (hay independencia entre filas y columnas), pues el valor obtenido es improbable para una tabla con filas y columnas independientes. En este caso, suponemos que las variables están asociadas.

- Si el valor calculado χ_{exp}^2 tiene una probabilidad igual o mayor que α (nivel de significación) no podemos rechazar la hipótesis nula H_0 . En este caso no tomamos ninguna decisión.
- Una alumna pregunta como comparar el valor del estadístico con el valor de la tabla de la distribución Chi. El profesor se lo aclara.

El profesor pasa a explicar el contraste de homogeneidad, resaltando el hecho que es muy parecido al anterior, para que los alumnos tomen confianza con el tema. Indica que la principal diferencia con el anterior que son las hipótesis, y utiliza el ejemplo 3.3 sobre grupo sanguíneo para indicar los pasos necesarios en el contraste.

Se indica que se aplica al caso en que la tabla de contingencia presenta una población X clasificada en r subpoblaciones x_1, x_2, \dots, x_r . En cada una de estas poblaciones se toma una muestra, y los individuos de la misma se clasifican según una variable Y que puede tomar m valores posibles y_1, y_2, \dots, y_m . Sea p_{ij} la proporción de individuos que, en la población x_i tiene como valor de $Y=y_j$.

En el *contraste de homogeneidad* es cuando se desean contrastar las dos hipótesis siguientes:

- $H_0: p_{1j} = p_{2j} = \dots = p_{mj}$ para todo j ; dicho de otro modo, todas las subpoblaciones tienen idéntica distribución para la variable Y .
- H_1 : algunas de estas proporciones son diferentes. Dicho de otro modo, la distribución de la variable Y en alguna de estas subpoblaciones es diferente

Se desarrollan los mismos pasos que en el caso del contraste de independencia, calculando las frecuencias esperadas y el estadístico Chi-cuadrado en el ejemplo, obteniendo $\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 1,76$ con 3 grados de libertad.

De la tabla Chi-cuadrado obtenemos que la probabilidad de obtener un valor 7,81 o mayor con 3 grado de libertad es $p = 0,184$. Por tanto el valor es no estadísticamente significativo, pues es mayor que 0,01. No se puede rechazar la hipótesis de homogeneidad de grupos sanguíneos en las dos muestras, por lo que suponemos las muestras son homogéneas.

Para concluir el tema se estudian las condiciones de aplicación del chi cuadrado. El profesor comenta el hecho de que no se puede aplicar siempre este estadístico, pasando a mostrar los casos en que se puede aplicar. Observa que al estudiar el valor de Chi-cuadrado en la tabla de la distribución, obtenemos siempre un valor positivo. Es decir, siempre hacemos un contraste unilateral.

También se indica que, si las frecuencias esperadas en las celdas son muy pequeñas, puesto que en la fórmula $\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$ aparecen dividiendo, se obtendría un valor alto de Chi-cuadrado, aunque las diferencias entre frecuencias observadas y esperadas fuesen grandes. Por eso, se recomienda que se use una muestra de suficiente tamaño. Indica que dos recomendaciones importantes son: (a) Como máximo el 20% de las frecuencias esperadas pueden ser menores que el valor 5 y (b) No debe usarse si hay frecuencias esperadas inferiores a 1.

Después de finalizar el tema se hace un resumen de lo estudiado Los alumnos no hacen preguntas durante este tiempo.

A continuación el profesor propone resolver el Ejemplo 3.2 (sobre estudio de asociación entre asociación entre la práctica deportiva y la sensación de bienestar), para lo cual deja un poco de tiempo y recuerda que lo primero es calcular las frecuencias esperadas, poniendo la fórmula de estas frecuencias en la pizarra.

- El profesor dando vueltas por la clase, y se acerca a un alumno que está muy parado y parece no saber cómo comenzar el ejercicio. El profesor se da cuenta que no sabe calcular las frecuencias esperadas y se lo explica de forma individual.
- El profesor se acerca a otro alumno que ya ha sido capaz de calcular las frecuencias esperadas y también ha calculado el valor del estadístico Chi-cuadrado y los grados de libertad, pero tiene dificultades para encontrar el valor crítico en la tabla. El profesor recuerda a los estudiantes cómo obtenerlo.
- Como algunos estudiantes están teniendo problemas en el cálculo de las frecuencias esperadas, el profesor pregunta en voz alta si alguien se ofrece para decírselas. Al no ofrecerse nadie el profesor pregunta a varios alumnos valores que dicen correctamente. Por tanto, no parece haber más problemas en este punto, sólo timidez para ofrecerse voluntario para responder.

El profesor vuelve a dejar un poco de tiempo, dando indicaciones de cómo seguir el problema una vez calculadas las frecuencias esperadas.

- Un alumno pregunta por el significado p-valor, y el profesor dice que ya se ha explicado anteriormente cuando se vieron otros contrastes. Pero que de todas formas les comenta lo que viene explicado en los apuntes y la página donde pueden repasarlo. También les indica que no es necesario para este ejercicio, ya que se puede resolver el contraste comparando los valores del estadístico con el valor crítico de la tabla. Recuerda cómo hacer la comparación e indica que viene explicado en los apuntes.

El profesor mientras los alumnos trabajan va dando vueltas por la clase acercándose a los grupos donde están discutiendo algunos pasos entre sí, para aclararles posibles dudas.

- El profesor va preguntando a algunos alumnos los valores de la tabla de componentes de chi-cuadrado, los cuales dan los alumnos correctamente, preguntando finalmente el valor del estadístico Chi cuadrado, que también se ha calculado correctamente.
- El profesor pregunta los grados de libertad y una alumna los da correctamente.
- Una alumna pregunta sobre la diferencia entre asociación e independencia, y se le aclara.

El profesor va mostrando los resultados conforme los grupos de alumnos los finalizan correctamente. Se finaliza e interpreta correctamente el contraste de independencia..

A continuación el profesor propone resolver el Ejercicio 3.1 (inmigración), para lo cual deja un poco de tiempo y recuerda que lo primero es calcular las frecuencias esperadas, poniendo otra vez la fórmula de estas frecuencias en la pizarra.

- El profesor pregunta las frecuencias esperadas y las van diciendo los alumnos en forma correcta. No parece haber problemas con el significado y cálculo de estas frecuencias.

- Una alumna pregunta “¿por qué alfa vale 0,01?”, no se da cuenta que viene en el enunciado.
- Los alumnos siguen dando los resultados de los cálculos de la tabla de componentes de Chi cuadrado y del estadístico Chi cuadrado, los grados de libertad y el valor crítico, hasta que los dan todos de forma correcta.
- El profesor enfatiza que hay que interpretar los resultados de los problemas en su contexto y realiza una interpretación de los resultados en este ejercicio.

El profesor va mostrando los resultados conforme lo van diciendo bien. Este día se pudo finalizar todo el tema. El profesor pide a los estudiantes que lean con detalle los apuntes en su casa.

Sesión 4. Tema 4. Medidas de asociación

Se recuerda a grandes rasgos lo visto el día anterior, entablando una relación con el siguiente tema. Se les pasa la lista para controlar la asistencia y se les recuerda la existencia de la página web, con los materiales y el correo electrónico del profesor para hacer preguntas. Se introducen los objetivos del tema, que son interpretar la intensidad de la dependencia entre dos variables en una tabla de contingencia: calcular e interpretar medidas de asociación en tablas 2x2: y tablas rxc, diferenciando las que informan de la reducción del error de predicción de una variable, cuando se conoce el valor de la otra.

Se comienza recordando el test chi-cuadrado, que sirve para ver si dos variables están o no asociadas. Se indica que este contraste sólo nos informa de la presencia de asociación, pero no nos dice si es alta o baja. Por tanto, si se rechaza la hipótesis de independencia, el siguiente paso sería calcular algún valor que mida la *intensidad de la asociación*, es decir, indique cómo de dependientes son las variables la una de la otra.

Continúa la explicación diciendo que se va a mostrar distintos coeficientes que sirven para medir esta intensidad y que se diferenciará el caso de las tablas 2x2 de las tablas rxc.

Se indica que se van a ver primero las medidas de asociación para tablas 2x2. Se recuerda que en estas tablas, además de la intensidad de la asociación podemos ver el signo y mediante la tabla 4.1 se muestra la forma de las mismas, que muestran la presencia o ausencia de dos factores A y B.

Se explica primero el coeficiente Phi de Pearson, basado en el valor chi-cuadrado. Se recuerda que para las tablas 2x2 podemos utilizar la forma alternativa para el cálculo de Chi-cuadrado: $\chi^2_{\text{exp}} = \frac{(f_{11}f_{22} - f_{12}f_{21})^2 n}{f_{1\cdot} \cdot f_{2\cdot} \cdot f_{\cdot 1} \cdot f_{\cdot 2}}$ Y a partir de ahí, se define el coeficiente Phi, de la forma siguiente:

$$\Phi = \sqrt{\chi^2 / n} = \sqrt{\frac{(f_{11}f_{22} - f_{12}f_{21})^2}{f_{1\cdot} \cdot f_{2\cdot} \cdot f_{\cdot 1} \cdot f_{\cdot 2}}}$$

Donde se da su formula, el rango y su interpretación, indicando que el coeficiente toma valores entre -1 y 1:

- El valor máximo (1) se obtiene cuando la dependencia es directa y perfecta, todos los casos están en las celdas f_{11} y f_{22} . Si el coeficiente es positivo, la dependencia es directa y más alta cuanto más se acerque a 1.

- El valor mínimo (-1) se obtiene cuando la dependencia es inversa y perfecta, todos los casos están en las celdas f_{12} y f_{21} . Si el coeficiente es negativo, la dependencia es inversa y más alta cuanto más se acerque a -1.
- El valor 0 se obtiene cuando hay independencia.

Para aclarar la explicación del coeficiente se aplica en el ejemplo 4.1 (observación de conducta de interacción en niños en situación de juego). Se calculan primero las frecuencias esperadas y a partir de ellas obtenemos el valor Chi-cuadrado, calculando finalmente el coeficiente Phi de Pearson: $\Phi = \sqrt{\chi^2 / n} = 0,675$

Se resalta que el valor es positivo y por tanto indica dependencia directa. Se hace ver a los alumnos que, en efecto, aparecen muchos más datos en la diagonal principal f_{11} y f_{22} que en la otra diagonal. También se hace notar que el valor es moderado-alto (cerca de 0,7).

Para mostrar la utilidad de coeficiente se indica que si tuviéramos un valor del observador A, sabiendo el tipo de resultado, podríamos tratar de predecir su resultado por el observador B. La mayoría de los que tienen un resultado A por el observador A, también obtienen un resultado A por el observador B (e igual con los resultados de B). Cuantos más casos en la diagonal principal, mayor valor de Phi y más fácil la predicción.

Puesto que algunos alumnos tienen dificultad con el signo de la raíz cuadrada, el profesor da indicaciones sobre cuando se coge el signo positivo y cuando el negativo de la raíz cuadrada de la formula.

Se continúa explicando el riesgo relativo por filas y por columnas. Se explica su formula, el rango y su interpretación. Se hace ver que el *riesgo relativo por columnas* indica cuanto más probable es la presencia de A con B que entre aquellos que no poseen B. El *riesgo relativo por filas* indica cuanto más probable es la presencia de B con A que entre aquellos que no poseen A.

Se presentan las fórmulas de los dos coeficientes y se hace notar que estos dos valores puede que coincidan, pero esto no pasa siempre. Finalmente se interpreta el valor del Riesgo relativo cambia según el tipo de asociación que tengan las variables:

- El $RR = 1$, informa que no hay asociación entre las variables.
- El $RR > 1$, nos dice que existe asociación positiva.
- El $RR < 1$, indica que existe una asociación negativa.

Para aclarar lo explicado se aplica en el mismo ejemplo 4.1 (observación de conducta). Se obtienen los siguientes valores:

$$RR_{columnas} = \frac{P(ObA_A/ObB_A)}{P(ObA_A/ObB_B)} = \frac{100/120}{10/70} = 5,8333$$

Al ser el $RR_{columnas} > 1$, indica asociación positiva. El coeficiente nos dice que es 5,8333 veces más fácil tener un valor A por el observador A cuando se tiene un valor A por el observador B que si se tiene un valor B por el observador B.

$$RR_{filas} = \frac{P(ObB_A/ObA_A)}{P(ObB_A/ObA_B)} = \frac{100/110}{20/80} = 3,6364$$

El $RR_{filas} > 1$, nos dice que existe asociación positiva. El coeficiente indica que es 3,6364 veces más fácil tener un valor A por el observador B cuando se tiene un valor A por el observador A que si se tiene un valor B por el observador A.

- Un alumno dice que no comprender la interpretación en el ejemplo. El profesor vuelve a repetir la interpretación en la pizarra. Parece que el ejemplo al ser las dos variables parecidas genera problemas de comprensión de la diferencia entre el riesgo por filas y columnas..

Se explica el último coeficiente para tablas 2x2, siendo la razón de productos cruzados. Este coeficiente es una razón de dos cocientes: $RC = \frac{C_1}{C_2}$, donde C_1 es la razón de casos en que se presenta A y los que no se presenta A cuando está presente B y C_2 es la razón de casos A y no A cuando no está presente el factor B .

Se hace observar que la razón de productos cruzados es una medida no simétrica. Es decir, A es la variable dependiente y B la independiente. Se comenta la interpretación del coeficiente:

- Si $RC = 1$, hay la misma razón de casos que aparece A y \bar{A} , cuando está B , que cuando no está presente B ,
- Cuando $RC < 1$, la razón entre los casos que aparecen A y \bar{A} es menor cuando está presente B .
- Si $RC > 1$, la razón entre los casos que aparecen A y \bar{A} es mayor cuando está presente B .

Una vez explicada su fórmula el rango y su interpretación, para aclararlo se aplica en el mismo ejemplo 4.1 (observación de conducta).

- Una alumna pregunta sobre la interpretación del coeficiente en el ejemplo 4.1, antes de explicarla el profesor. El profesor la explica con la diapositiva, que, ya que en el ejemplo se obtiene $RC = \frac{f_{11}f_{22}}{f_{12}f_{21}} = \frac{100/20}{10/60} = \frac{100 \times 60}{10 \times 20} = 30$. Esto indica que, entre los sujetos con resultados

A del observador B hay 100 valorados con A por el observador A por cada 20 valorados con B por el observador B (la razón es 5/1). Entre los sujetos con resultados B del observador B hay 10 valorados con A por el observador A por cada 60 valorados con B por el observador B (la razón es 1/6). Por tanto hay relación entre el observador y el tipo de observación que realiza. El alumno sigue sin entender la explicación y el profesor vuelve a explicarlo utilizando la pizarra. Como hemos dicho, en este ejemplo, se debiera haber cambiado el símbolo utilizado en una de las variables, pues se prestó a confusión.

- Un alumno pregunta que es la razón, el profesor lo explica para el caso general y lo aplica al ejemplo.
- Para reforzar la explicación, el profesor señala los valores que corresponden a la asociación directa y cual a la asociación inversa, lo cual parece tranquilizar a los alumnos que toman notas. Estos alumnos, con no mucho bagaje matemático prefiere en general aprender algoritmos sencillos que puedan aplicar a la situación.

El profesor indica que a continuación se van a estudiar los coeficientes de asociación para tablas rxc. En estas tablas ya no se puede diferenciar el signo de la asociación, sino tan sólo su intensidad.

Se comienza explicando el coeficiente de contingencia de Pearson, donde se explica su fórmula, $C = \sqrt{\chi^2 / (\chi^2 + n)}$, el rango de valores que toma, desde 0 hasta un máximo Max

$\{C\} = \frac{\sqrt{\text{Min}\{r-1, c-1\}}}{\sqrt{1 + \text{Min}\{r-1, c-1\}}}$ y su interpretación. Este coeficiente cuando vale 0 indica independencia absoluta, y a mayor valor es mayor la asociación.

Para aclarar su significado, se aplica en el ejemplo 4.2 (relación entre actitud hacia el aborto y estado civil). Se calculan las frecuencias esperadas (tabla 4.2) y se hace observar que hay más solteros con actitud abortista de lo esperado (y menos contrarios al aborto); hay menos casados con esta actitud que lo esperado y algunos divorciados más. Por tanto hay relación entre las variables.

Al calcular el valor Chi-cuadrado, $\chi^2_{\text{exp}} = 145,83$ se obtiene un valor estadísticamente significativo (se puede comprobar mirando la tabla para 2 grados de libertad. A partir de él se obtiene el coeficiente $C = \frac{\sqrt{\chi^2}}{\sqrt{\chi^2 + n}} = \frac{\sqrt{145,83}}{\sqrt{145,83 + 500}} = 0,475$

Para interpretarlo se comenta que en este caso hay que calcular también el valor máximo, que sería $\text{Max } \{C\} = \frac{\sqrt{\text{Min}\{r-1, c-1\}}}{\sqrt{1 + \text{Min}\{r-1, c-1\}}} = 0,7071$. Por ello, podemos decir que la asociación es moderada-alta.

- Una alumna no entiende cómo se obtiene el máximo. El profesor se lo explica.

Se continua explicando el coeficiente V de Cramer, donde se explica su formula, $V = \frac{\sqrt{\chi^2 / n(p-1)}}{\sqrt{p-1}}$ siendo $p = \text{Min}\{\text{número de filas, número de columnas}\}$. Se estudia el rango y su interpretación. Este coeficiente varía entre 0 y 1, siendo 0 en caso de independencia y 1 en caso de dependencia perfecta.

Para aclararlo se aplica en el ejemplo 4.2 (sobre actitud hacia el aborto). Para el ejemplo dado El valor chi-cuadrado era 145,83. El coeficiente se calcula mediante la siguiente formula, siendo $p = \text{Min}\{\text{número de filas, número de columnas}\}$:

$$V = \frac{\sqrt{\chi^2 / n(p-1)}}{\sqrt{p-1}} = \frac{\sqrt{145,83 / 500 \times (2-1)}}{\sqrt{2-1}} = \sqrt{0,29166} = 0,54$$

Como los coeficientes anteriores, el valor obtenido indica una asociación moderada.

Se pasa a las medidas basadas en la reducción proporcional del error. Para motivar su introducción se hace el siguiente supuesto: Supongamos que quiero predecir el valor de la característica X (variable en filas) en un individuo tomado al azar en la población.

- Si no tuviera ninguna información sobre el mismo, la probabilidad de cometer un error en la clasificación sería: $P(\varepsilon_1) = \frac{n - f_{\text{max}}}{n}$, siendo f_{max} la mayor frecuencia marginal en filas. Ya que si no se nada, lo asignaría a la fila de mayor frecuencia.
- Si supiera cuál es la columna (valor de la variable Y), lo asignaría a la fila (valor de X) que tenga frecuencia máxima en esa columna. La probabilidad de error sería $P(\varepsilon_2) = \frac{n - \sum f_{mj}}{n}$ donde f_{mj} es la frecuencia máxima en la columna j.
- El coeficiente $\lambda = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1}$ indica cual es el porcentaje de error que se ve reducido al predecir el valor de la variable dependiente X, conocido el valor de la variable independiente Y, en vez de asignar al azar el valor de X.

Se continúa explicando el desarrollo de la fórmula del coeficiente Lambda de Goodman y Kruskal, el rango y su interpretación. Para facilitar su comprensión se aplica en el ejemplo 4.2 (actitud hacia el aborto), haciendo las siguientes consideraciones:

- Supongamos que queremos predecir el estado civil de una persona al azar en esta muestra, sin saber nada sobre su actitud respecto al aborto. Como la máxima frecuencia marginal en filas es la correspondiente a casados (250), diríamos que es casado. La probabilidad de error en este caso sería $P(\varepsilon_1) = \frac{n - f_{\max}}{n} = \frac{500 - 250}{500} = 1/2$; pues nos equivocaríamos con los solteros y divorciados.
- Si nos dicen cuál es la actitud de la persona, predeciríamos que es soltero (si tiene actitud abortista) o casado (si no la tiene). La probabilidad de error ahora sería $P(\varepsilon_2) = \frac{n - \sum f_{mj}}{n} = \frac{500 - (120 + 200)}{500} = \frac{180}{500} = 0,36$. Hemos reducido el error de predicción, ahora sólo nos equivocamos en el 36% de los casos, en vez de la mitad
 - El cociente $\lambda = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1} = \frac{0,5 - 0,36}{0,5} = \frac{0,14}{0,5} = 0,28$

Esto indica que hemos reducido un 28% los errores de predicción de la actitud hacia el aborto, si sabemos el estado civil de la persona.

- Un alumno indica que no entiende como se obtiene el desarrollo de la fórmula. El profesor vuelve a explicar la definición del coeficiente y el modo en que se obtiene el desarrollo.

Después de finalizar el tema se hace un resumen de todo lo tratado, recordando la diferencia entre cada coeficiente, los valores que toman y su interpretación. Los alumnos no hacen preguntas durante este tiempo. Aparentemente el tema ha sido comprendido. A continuación el profesor propone resolver el Ejercicio 4.1 (calificaciones de estadística de cuatro profesores). En dicho ejercicio se pide calcular los coeficientes de contingencia de Pearson V de Cramer y Lambda de Goodman y Kruskal (con filas dependientes) e interpretarlos.

El profesor va mostrando los pasos y resultados que se van teniendo. Calcula primeramente las frecuencias esperadas y a partir de ellas el valor Chi-cuadrado $\chi_{\text{exp}}^2 = 23,80$. Los grados de libertad son 3. Para 3 grados de libertad, se busca en la tabla el valor crítico y se ve que el valor 12,84 tiene una probabilidad menor que 0,005. Por tanto el contraste es estadísticamente significativo y hay asociación entre profesor y calificación.

El coeficiente de contingencia de Pearson sería $C = \sqrt{\chi^2 / (\chi^2 + n)} = 0,208$. El valor máximo en este caso es: $\text{Max } \{C\} = 0,7071$. Por tanto la intensidad de la asociación es pequeña pues sólo llegamos a la tercera parte del máximo.

El coeficiente V de Cramer es: $V = \sqrt{\chi^2 / n(p-1)} = 0,213$. Como sabemos este coeficiente puede llegar a 1, por lo que confirma que la asociación es pequeña

La lambda de Goodman y Kruskal: $\lambda_x = \frac{(\sum f_{mj}) - f_{m+}}{N - f_{m+}} = 0,08$, de lo que observamos que no mejora mucho la predicción al saber si el alumno ha aprobado para predecir su profesor.

El profesor hace observar que el estudio conjunto de los tres coeficientes indica que todos los coeficientes están por debajo de la mitad del rango que pueden tomar, sin llegar a ser nula. Por lo que, aunque la calificación depende del profesor, la asociación es baja.

Finalmente por ser el último día de teoría el profesor pregunta si hay dudas de cualquier tema, nadie pregunta nada. El profesor se pone a comentar que les pondrá unas preguntas tipo verdadero o falso, para que practiquen de cara al examen final. También comenta que pondrá cuatro preguntas del examen final.

- Un alumno pregunta si será una de cada tema, el profesor no quiere darles pistas para que estudien a fondo los cuatro temas, pero que en una pregunta de un tema pueden venir cosas de temas anteriores.

Este día se pudo finalizar todo el tema. El profesor pide a los estudiantes que lean con detalle los apuntes en su casa.

Sesión 5. Actividades prácticas

El profesor comienza comentando como va a ser el funcionamiento de las prácticas, mientras se distribuyen los ejercicios y la copia del programa Excel. El profesor dice que la asistencia se controlará y que en la pauta que se ha entregado para recoger en forma escrita los ejercicios hay que poner el nombre.

Además les pide que intenten resolver los ejercicios por si mismos, una vez se les explique el funcionamiento del programa. Se les recuerda que a todos se va a considerar la participación, si muestran interés, aunque tengan algunos fallos. El objetivo de la práctica es que practiquen para los exámenes, Por ello, deben intentarlo de forma individual y no se preocuparse si tienen algún error, ya que no les cuenta.

Se les comunica que el material esta en la pagina web, y que las soluciones de las prácticas se colgaran también la siguiente semana. Algunos alumnos háganlo intentan solos, pero otros preguntan al que tienen al lado, aunque se les insiste en que no se preocupen por los fallos y que se les dará las soluciones en Internet. Algunos piden ayuda en la resolución, de modo que habrá que tener en cuenta esta forma de trabajo al interpretar los resultados.

Al comienzo de la clase el profesor les informa como funciona el programa, mostrándoles las diversas opciones y pantallas. Para la práctica 1 se utilizan las hojas 1 (Frecuencias), 2 (gráficos) y de la hoja 3 sólo se necesita la tabla que proporciona las frecuencias esperadas en caso de independencia.

Se les explica que el programa está protegido y por tanto, sólo deben rellenar las celdas en color verde, que corresponden a las celdas de datos y el resto son cálculos o gráficos proporcionados por el programa. Para la práctica 1 sólo se requiere dar la tabla de contingencia (en la hoja 1) como dato, pues todos los demás cálculos y gráficos los proporciona el programa.

El profesor comenta que cuando lo pongan, tendrán un ejemplo relleno que se ve en teoría (supervivencia en el Titanic), comenta que tienen que adaptarlo a cada ejercicio que se les propone, y responder a las preguntas observando los resultados.

El profesor le comenta a un alumno que llega tarde todo de nuevo. El profesor al ver que los alumnos van justos de tiempo dice que dejen para el final los gráficos pedidos en el problema 2, pues lleva más tiempo para copiarlo

A continuación se clasifican las dudas expresadas por los estudiantes para el funcionamiento general del programa y para cada uno de los ejercicios propuestos.

Preguntas referidas al funcionamiento programa:

- Un alumno pregunta si, para resolver los problemas hay que cambiar los datos que vienen. El profesor aclara que si.
- Otro alumnos preguntan si se pueden cambiar los datos. Del ejemplo incluido en el programa. El profesor indica que sí, pero sólo en las celdas de color verde, pues el resto está protegida.
- El profesor se pone a dar vueltas y va mirando si introducen bien los datos, muchos introducen la fila y la columna del total como si fuera otra variable, el profesor comprueba que todos lo tienen bien.

Dificultades relacionadas con el primer ejercicio

Actividad 1. En un colegio se pregunta a los alumnos, obteniendo los siguientes resultados:

	Chicos	Chicas	Total
Le gusta el tenis	400	200	600
No le gusta	50	50	100
Total	450	250	700

Si elegimos al azar uno de estos alumnos:

- a. ¿Cuál es la probabilidad de que le guste el tenis?
- b. ¿Cuál es la probabilidad de que sea una chica y además le guste el tenis?
- c. Sabiendo que el alumno elegido es chica, ¿Cuál es la probabilidad de que le guste el tenis?

En total se hacen 13 preguntas, que se pueden clasificar en las siguientes, ya que alguna la hizo más de un alumno:

- Un alumno indica que no entiende lo que hay que hacer. El profesor le ayuda a introducir los datos del primer ejercicio en el programa.
- Algunos alumnos lo resuelven correctamente, pero quieren confirmarlo enseñándoselo al profesor.
- Un alumno indica que no recuerda las probabilidades, y dice que suspendió la primera asignatura.
- Un alumno pregunta si se pueden intercambiar los valores de filas y columnas. El profesor indica que si, pues no afecta a los resultados.
- Algunos alumnos tienen dificultad para decidir si en la variable sexo las columnas de chicos y chicas debieran intercambiarse de orden.
- Un alumno pregunta si el apartado b y c del problema no preguntan lo mismo, por tanto está confundiendo la probabilidad condicional y conjunta. El profesor se lo aclara.

- Otro pregunta si en la fórmula del cálculo de la probabilidad condicional hay que dividir el numerador por la probabilidad del condicionante o del condicionado. El profesor recuerda la fórmula e indica que el problema se puede también resolver directamente a partir de los datos como cociente entre casos favorables y posibles.
- Un alumno se deja influir por sus teorías previas, dando su opinión sin usar los datos.

Dificultades relacionadas con el ejercicio 2

Actividad 2. Un psicólogo en un estudio obtuvo la siguiente clasificación de los pacientes de un hospital de acuerdo con las variables “sexo” y “diagnóstico”.

Sexo	Diagnóstico	
	Esquizofr.	No esquizofr.
Varón	43	15
Mujer	32	52

- ¿Hay asociación entre las variables? ¿Es directa o inversa?
- ¿Por qué?

En total se hacen 21 preguntas, todas en torno a las tres cuestiones siguientes

- No se entiende el ejercicio. Los alumnos quieren decir que no saben como resolverlo. Alguno había faltado a la clase teórica donde se explicó y otros no recordaban cómo se puede evaluar la asociación a partir de la tabla 2x2.
- ¿Es lo mismo asociación qué dependencia? Se recuerda a los alumnos que asociación es un concepto algo más amplio, mientras que dependencia sería un caso de la asociación. Se recuerda la diferencia entre los dos conceptos y se indica que coloquialmente dependencia se usa como sinónimo de asociación.
- No saben si la dependencia es directa o inversa, puesto que la variable sexo no es ordinal. Los alumnos tenían razón en este punto, puesto que aunque la tabla muestra, matemáticamente, dependencia directa, al ser la variable no ordinal se podrían intercambiar las filas y aparecería dependencia inversa. Se optó por suprimir la parte en que se pedía el signo. Esta fue la pregunta más frecuente.

Dificultades relacionadas con el ejercicio 3

Actividad 3. Para comprobar si existía relación entre el tabaco y la hipertensión, fueron elegidas aleatoriamente 100 personas, de las que 55 resultaron ser hipertensas.

Agrupadas dichas personas en tres clases, según que no fumaran, fumaran moderadamente o fumaran mucho, se encontró que entre los hipertensos 10 no fumaban y 25 lo hacían moderadamente. En cambio, entre los no hipertensos 18 fumaban moderadamente y 7 fumaban mucho.

- Organiza los datos para formar una tabla de doble entrada o de contingencia.
- Calcula las distribuciones condicionales por filas.
- Dibuja el diagrama de barras adosado y diagrama de barras apilado.

En total se hacen 1 preguntas, que se pueden clasificar en las siguientes, ya que alguna la hizo más de un alumno:

- El profesor recomienda que para construir la tabla, lo hagan a mano.
- Un alumno pregunta si en el ejercicio de construir la tabla, hay que copiarla en el papel. El profesor indica que si, para saber que está bien resuelto.
- El profesor ve un alumno dibujando los gráficos a mano, y le recuerda que en la segunda pestaña del programa aparecen y no necesita realizarlos a mano.
- Algunos alumnos preguntan qué son las distribuciones condicionales (pues en los apuntes se hablaba de frecuencias condicionales sin usar la palabra distribución). Se aclaró este punto que fue uno de los que más confundieron a los alumnos.
- Alguno preguntó si las distribuciones condicionales eran la suma de las frecuencias. El profesor razonó que entonces obtendría la unidad.

Dificultades en el cuarto ejercicio

Actividad 4. Un psicólogo en un estudio obtuvo la siguiente clasificación de una muestra de 100 sujetos de acuerdo con las variables, “concepción de la inteligencia” y “tipo de autoinstrucciones”

Concepción	Autoinstrucción		
	Respuestas Instruccionales	Respuestas atribucionales	Otras respuestas
Destreza	24	11	9
Rasgo	72	33	27

Comprueba si se cumplen las propiedades de independencia.

En total se hacen 18 preguntas, que se pueden clasificar en las siguientes, ya que alguna la hizo más de un alumno:

- Los alumnos encuentran una errata en el enunciado del ejercicio 4, ya que no son 100 sujetos.
- Un alumno piensa que no se cumple una propiedad de independencia, pero al repetir las cuentas se da cuenta de que el programa redondea.
- Otro pregunta si se comprueba la independencia multiplicando (No saben interpretar la igualdad $h_{ij} = h_{i.} \times h_{.j}$ debido a que no interpretan los símbolos). Se les recuerda el procedimiento y se explica la notación.
- Se vuelve a repetir la pregunta del cálculo de distribuciones condicionales (posiblemente este alumno no prestó atención en el ejercicio anterior a la explicación).
- Un alumno comprueba dos propiedades de la independencia (igualdad de distribuciones condicionales y con la marginal) y la igualdad de frecuencias observadas y esperadas. Pregunta si hay más propiedades y se le recuerda que $h_{ij} = h_{i.} \times h_{.j}$
- Otro alumno indica que le falta una propiedad por comprobar. El profesor mira las que ha mostrado y ve que le falta comprobar la igualdad de frecuencias observadas y esperadas.
- Algunos alumnos confunden la notación h y f (no saben si se refiere a frecuencias absolutas o relativas) con lo cual hay que explicarlo.




ANEXO 5
DESARROLLO TEÓRICO DEL TEMA



Tema 1. Tablas de contingencia: Lectura e interpretación

Objetivos del tema 1

<p>En este tema aprenderás a:</p> <ul style="list-style-type: none">• Resumir datos sobre dos variables estadística en una tabla de contingencia.• Identificar las frecuencias que corresponde a cada par de valores de las variables.• Calcular las frecuencias relativas dobles e interpretarlas.• Calculas las frecuencias marginales, e interpretarlas.• Calcular las frecuencias condicionales e interpretarlas.• Representar gráficamente los datos mediante diagrama de barras adosadas, diagrama de barras apiladas y gráfico tridimensional.• Calcular probabilidades simples, compuestas y condicionales a partir de datos de una tabla de contingencia.	<p>Primera aparición de tablas de contingencia</p>  <p>KARL PEARSON</p>
--	---

1.1. Introducción

Hasta este momento has estudiado las variables unidimensionales, en las cuales calculaste las medidas de posición central o medidas de dispersión y otros estadísticos. En dicho estudio las variables se presentaron de forma aislada, cosa que no es habitual en la vida real, ya que al estudiar el estado físico de una persona, por ejemplo, se realizan una gran variedad de preguntas como altura, peso, si realiza ejercicio, etc. Todas estas, y muchas más, se realizan por la sencilla razón de que en multitud de ocasiones las variables están interrelacionadas entre sí. Siguiendo con el ejemplo de estado físico, una persona alta es razonable suponer que tiene mayor peso, dos variables comúnmente relacionadas.

Recuerda: En un estudio estadístico podemos encontrar *variables cuantitativas* (como son: edad, peso, altura...) y *variables cualitativas* (como son: color de ojos, color de pelo...)

En este tema vamos a estudiar las posibles relaciones entre dos variables cualitativas, mediante una herramienta conocida como tabla de contingencia.

1.2. Tablas de contingencia

La tabla de contingencia nos proporciona una forma resumida de representar datos, de dos variables que se quieren estudiar.

Ejemplo 1.1. Problemas neuronales

En un hospital psiquiátrico se hace un estudio en el que participan 30 pacientes con dos tipos de problemas neuronales (altos y bajos), queremos comparar un fármaco nuevo con otro antiguo. ¿Cómo podemos representar esta situación? ¿Cómo podemos ver si el tratamiento nuevo es preferible al anterior?

Para representar esta situación utilizaremos dos variables. La **variable X** representa el tratamiento, con dos valores (x_1, x_2). La **variable Y** los problemas neuronales (y_1, y_2). Los datos se representan en una tabla como la 1.1 que llamaremos *tabla de contingencia*. En este ejemplo mostramos el caso más

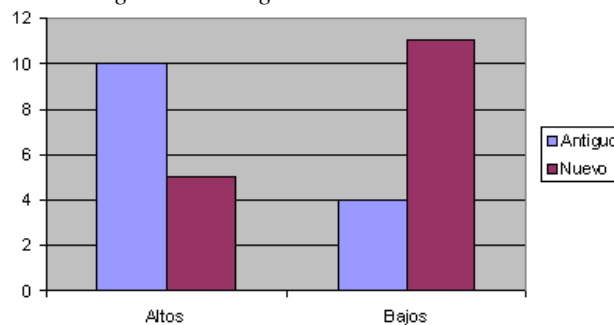
sencillo de tabla de contingencia, ya que las variables en filas y columnas solo tienen dos valores cada uno. Se trataría de una tabla 2x2 (también llamada dicotómica).

Una posible representación de esta tabla sería el diagrama de barras adosadas (Figura 1.1.). En este diagrama usamos barras colocadas horizontalmente o verticalmente para representar la frecuencia de cada casilla de la tabla de contingencias. Para cada valor de la variable X (tratamiento antiguo o nuevo) representamos una al lado de otra la frecuencia con que se presenta cada valor de Y (problemas neuronales altos o bajos) dentro de ese valor de X. También se podría dar el diagrama de barras en porcentajes, bien absolutos (respecto al total de la tabla), bien relativos (respecto a cada valor de la variable X).

Tabla 1.1. Tipo de tratamiento en enfermos con problemas neuronales

Tipo de tratamiento(X)	Problemas neuronales (Y)	
	Altos(y_1)	Bajos(y_2)
Antiguo(x_1)	10	4
Nuevo(x_2)	5	11

Figura 1.1. Diagrama de barras adosadas



1.3. Frecuencias dobles absolutas y relativas.

En la tabla 1.1 tenemos dos tipos de pacientes y dos tipos de tratamientos. Por tanto se presentan cuatro tipos de enfermos:

- Con problemas neuronales altos que usan el tratamiento antiguo (10 casos)
- Con problemas neuronales altos y que usan el tratamiento nuevo (5 casos)
- Con problemas neuronales bajos y que usan el tratamiento antiguo (4 casos)
- Con problemas neuronales bajos y que usan el tratamiento nuevo. (11 casos).

Para saber qué tipo es representado por cada valor en una celda de la tabla no hay más que seguir la fila y columna en la que está el valor. Por ejemplo, al seguir la columna correspondiente al valor 5, vemos que son pacientes con problemas neuronales altos, y al seguir la fila observamos que son tratados con el tratamiento nuevo, por tanto hay 5 pacientes con problemas neuronales altos que son tratados con el tratamiento nuevo. Llamamos frecuencias absolutas dobles a estos valores y los representamos mediante f_{ij} indica la frecuencia absoluta con que aparece el par (x_i, y_j) donde el primer número representa el valor que acompaña a la variable X y el segundo número representa el valor que acompaña a la variable Y (Ver tabla 1.2).

Tabla 1.2. Frecuencias absolutas dobles

Tipo de tratamiento(X)	Problemas neuronales (Y)	
	Altos(y_1)	Bajos(y_2)
Antiguo(x_1)	10(f_{11})	4(f_{12})
Nuevo(x_2)	5(f_{21})	11(f_{22})

En algunos casos, puede ser interesante tener las frecuencias en tanto por uno. Estos valores llamamos *frecuencias relativas doble*. Si representamos mediante h_{ij} la frecuencia relativa del par de valores (x_i, y_j) , su valor se obtiene dividiendo cada celda por el total de pacientes n (en nuestro caso 30), es decir:

$$h_{ij} = f_{ij}/n$$

En el ejemplo: $h_{11} = f_{11}/n = 10/30 = 0,333$; $h_{12} = f_{12}/n = 4/30 = 0,133$ y del mismo modo se obtiene el resto de las frecuencias relativas dobles (tabla 1.3). A partir de la tabla 1.3. es fácil calcular *el porcentaje de pacientes en cada celda respecto al total de la muestra* simplemente multiplicando por cien cada celda. En nuestro caso tenemos:

- Hay 33,33% pacientes con problemas neuronales altos y se les aplica el tratamiento antiguo.
- Hay 16,67% pacientes con problemas neuronales altos y se les aplica el tratamiento nuevo.

Y del mismo modo se obtiene el resto de los porcentajes.

Tabla 1.3. Frecuencias relativas dobles

		Problemas neuronales (Y)	
		Altos(y_1)	Bajos(y_2)
Tipo de tratamiento(X)	Antiguo(x_1)	0,333(h_{11})	0,133(h_{12})
	Nuevo(x_2)	0,167(h_{21})	0,367(h_{22})

1.4. Distribuciones marginales

A partir de la tabla de contingencia (bidimensional) pueden obtenerse diferentes distribuciones de una variable (unidimensional).

Si en la tabla de frecuencias se suman las frecuencias por columnas, obtengo en cada columna j , el número de individuos $f_{.j}$ con un valor de la variable $Y=y_j$, independientemente del valor X . A la distribución así obtenida se le conoce como *distribución marginal* de la variable Y . De forma análoga podemos definir la *distribución marginal* de la variable X .

Ejemplo 1.1. Problemas neuronales (continuación)

- Siguiendo el ejemplo, podemos observar en la tabla 1.4. el total por filas o *distribución marginal de X*: 14 enfermos con tratamiento antiguo y 16 con tratamiento nuevo representados por $f_{1.}$ y $f_{2.}$, o sea $f_{11}+f_{12}=f_{1.}$ y $f_{21}+f_{22}=f_{2.}$.
- De la misma manera vemos como el total por columnas o *distribución marginal de Y* (15 enfermos con problemas altos y otros 15 con problemas bajos, representados por $f_{.1}$ y $f_{.2}$).
- Al sumar la columna del total por filas, o sumando el total por columnas sale el mismo resultado: 30 enfermos, o total de la muestra ($f_{1.}+f_{2.} = f_{.1}+f_{.2} = n$).

Tabla 1.4. Frecuencias absolutas de las distribuciones marginales

		Problemas neuronales (Y)		
		Altos(y_1)	Bajos(y_2)	TOTAL
Tipo de tratamiento(X)	Antiguo(x_1)	10(f_{11})	4(f_{12})	14($f_{1.}$)
	Nuevo(x_2)	5(f_{21})	11(f_{22})	16($f_{2.}$)
TOTAL		15($f_{.1}$)	15($f_{.2}$)	30(n)

- Dividiendo todas las frecuencias marginales por n se obtienen las *frecuencias relativas marginales* o también sumando por filas o columnas las frecuencias relativas dobles. Además la suma de las frecuencias relativas marginales en las filas o columnas es igual a 1.

Tabla 1.5. Frecuencias relativas de las distribuciones marginales

		Problemas neuronales (Y)		
		Altos(y_1)	Bajos(y_2)	TOTAL
Tipo de tratamiento(X)	Antiguo(x_1)	0,333(h_{11})	0,133(h_{12})	0,467($h_{1.}$)
	Nuevo(x_2)	0,167(h_{21})	0,367(h_{22})	0,533($h_{2.}$)
TOTAL		0,5($h_{.1}$)	0,5($h_{.2}$)	1

1.5. Frecuencias relativas condicionales

Es posible que queramos centrarnos en solamente una parte de los pacientes, por ejemplo en los pacientes con el tratamiento nuevo. Para este tipo de casos se calculan *distribuciones condicionales*. Si representamos por $h(x_i/y_j)$ la frecuencia relativa condicional del valor x_i entre los individuos que presentan el carácter y_j , obtenemos la siguiente relación:

$$h(x_i | y_j) = \frac{f_{ij}}{f_{.j}} = \frac{h_{ij}}{h_{.j}}$$

Podemos también obtener la frecuencia relativa de y_j condicionada por $x=x_i$ mediante la expresión:

$$h(y_j | x_i) = \frac{f_{ij}}{f_{i.}} = \frac{h_{ij}}{h_{i.}}$$

En el ejemplo, para obtener la distribución condicional de tipo de problema neuronal dentro de los pacientes con tratamiento nuevo (tabla 1.6) tendríamos:

$$h(\text{problemas altos/ tratamiento nuevo}) = h(y_1/X=x_2) = f_{21}/f_{2.} = 5/16 = 0,3125$$

$$h(\text{problemas bajos /tratamiento nuevo}) = h(y_2/X=x_2) = f_{22}/f_{2.} = 11/16 = 0,6875$$

Tabla 1.6. Distribución de problemas neuronales en enfermos con tratamiento nuevo

Tipo de tratamiento(X)	Problemas neuronales (Y)		
	Altos(y_1)	Bajos(y_2)	TOTAL
Nuevo(x_2)	0,3125 ($h(y_1/\text{nuevo})$)	0,6875 ($h(y_2/\text{nuevo})$)	1

A partir de la tabla 1.6 podemos responder preguntas como: ¿Hay el mismo porcentaje con problemas altos y bajos en el nuevo tratamiento? En forma similar calculamos la distribución del tipo de tratamiento condicionada que el paciente tiene problemas neuronales bajos (ver tabla 1.7.)

Tabla 1.7. Distribución de X condicionada a Y_2

		Problemas neuronales (Y)	
		Bajos(y_2)	
Tipo de tratamiento(X)	Antiguo(x_1)	0,2667 ($h(x_1/Y=y_2)$)	
	Nuevo(x_2)	0,7333 ($h(x_2/Y=y_2)$)	
TOTAL		1	

1.6. Tablas rxc (caso generalizado)

A veces podemos tener más de dos categorías en alguna de las variables. El formato general de una tabla rxc sería parecido al formato visto en tablas 2x2. Básicamente la diferencia es que se le van añadiendo filas o columnas, calculándose las distintas frecuencias en la forma indicada

Tabla 1.8. Tabla de contingencia para el caso general

		Variable Y			
		Y_1	...	Y_c	TOTAL
Variable X	X_1	f_{11}	...	f_{1c}	$f_{1.}$

	X_r	f_{r1}	...	f_{rc}	$f_{r.}$
TOTAL		$f_{.1}$...	$f_{.c}$	N

Siendo X e Y , las variables de interés, X_1, \dots, X_r , los distintos valores de la variable X ; Y_1, \dots, Y_c , los distintos valores de la variable Y y $f_{11}, \dots, f_{1c}, \dots, f_{r1}, \dots, f_{rc}$, las frecuencias absolutas conjuntas de los pares.

1.7. Representaciones gráficas

En el estudio de las tablas de contingencia hay una gran variedad de gráficos con los que mostrar datos. A continuación mostraremos algunos.

Diagrama de barras adosadas. Este sistema de representación nos muestra con barras colocadas horizontalmente o verticalmente la frecuencia de cada casilla del interior de la tabla de frecuencias (Figura 1.1.). Para cada valor de la variable X se presentan conjuntamente la frecuencia de cada valor de Y condicionado a este valor de X . Al estar situadas unas junto a otras permite la comparación rápida entre las variables y dentro de cada variable. También se podría dar el diagrama de barras en porcentajes, bien absolutos (respecto al total de la tabla), bien relativos (respecto a cada valor de la variable X).

Diagrama de barras apilado. Este gráfico muestra una barra por cada valor que tome la variable Y , las cuales a su vez, se dividen en distintos colores que representa a cada valor de la variable X . Representa la frecuencia con la que aparece cada valor de X en cada valor de Y , comparando entre categorías, el aporte de cada valor al total (Figura 1.2.).

En el ejemplo se da el diagrama de barras en porcentajes. De este modo se visualizan mejor las frecuencias condicionales que en el diagrama de barras adosado.

Figura 1.2. Diagrama de barras apilado

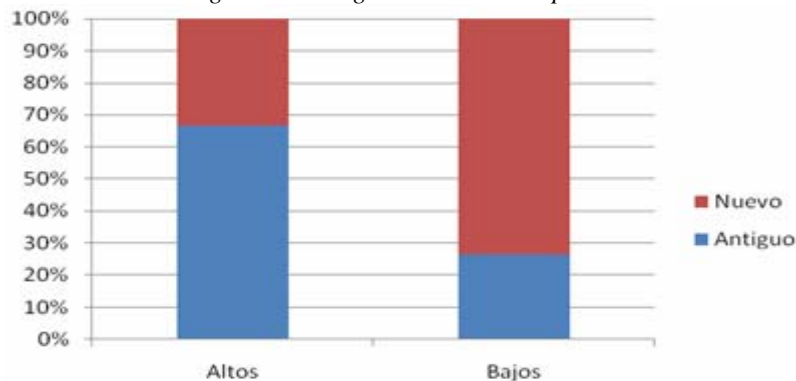
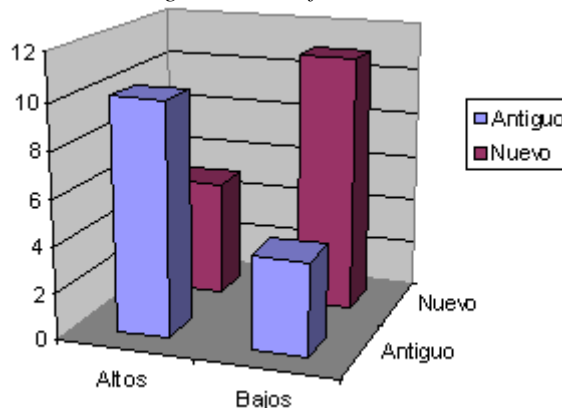


Gráfico tridimensional. Este gráfico tridimensional coloca en el ancho y largo las variables X e Y , y la altura representa la frecuencia de cada celda (Figura 1.3.). Se puede construir también utilizando frecuencias relativas o porcentajes.

Figura 1.3. Gráfico tridimensional



Ejercicios resueltos del tema 1

Ejercicio 1.1. Despersonalización

La despersonalización es una alteración de la percepción o la experiencia de uno mismo de tal manera que uno se siente "separado" de los procesos mentales o del cuerpo, como si uno fuese un observador externo a los mismos. La forma severa se encuentra en la ansiedad o en ataques de pánico. En la siguiente tabla clasificamos un grupo de pacientes según gravedad del síndrome de despersonalización y frecuencia de relaciones personales de la persona que lo presenta:

Relaciones personales (X)	Despersonalización (Y)	
	Bajo(y_1)	Alto (y_2)
Muchas(x_1)	24	22
Pocas(x_2)	19	40

1. Interpreta los datos de la tabla; calcula e interpreta las frecuencias marginales y relativas dobles y marginales.
2. Calcular e interpretar las frecuencias relativas condicionales por filas y columnas. Si se toma un enfermo al azar ¿Es más probable tener muchas relaciones personales en los pacientes con bajo o alto grado de despersonalización? ¿Es más probable tener alto grado de despersonalización en pacientes con pocas o muchas relaciones interpersonales?

1. Lectura de la tabla

Lo primero que tenemos que hacer para comprender la situación es leer la tabla, donde observamos las siguientes frecuencias absolutas dobles:

- Hay 24 personas con muchas relaciones personales y bajo nivel de despersonalización
- Hay 22 personas con muchas relaciones personales y alto nivel de despersonalización
- Hay 19 personas con pocas relaciones personales y bajo nivel de despersonalización
- Hay 40 personas con pocas relaciones personales y alto nivel de despersonalización

Frecuencias marginales

- Hay 46 personas con muchas relaciones personales (24+22) y 59 personas con pocas relaciones personales (19+40).
- Hay 43 personas con baja despersonalización (24+19) y 62 personas con alta despersonalización (22+40).
- Hay 105 personas en el experimento (24+22+19+40).

Frecuencias relativas dobles. Se calculan dividiendo cada frecuencia absoluta doble por el total de la muestra. Se presentan en la siguiente tabla

Relaciones personales (X)	Despersonalización (Y)	
	Bajo(y_1)	Alto (y_2)
Muchas(x_1)	0,2285	0,2095
Pocas(x_2)	0,181	0,381

- El 22,85% del total de personas tienen muchas relaciones y baja despersonalización (24/105).
- Igualmente se interpretan el resto de las frecuencias (Observamos que los cuatro valores anteriores suman 1)

Frecuencias relativas marginales. Se presentan en la tabla siguiente

Relaciones personales	Despersonalización			%
	Bajo	Alto	TOTAL	
Muchas	24	22	46	0,4381
Pocas	19	40	59	0,5619
Total	43	62	105	
	0,4095	0,5905		

- El 43,81% de los participantes tienen muchas relaciones personales (46/105).
- El 40,95% de los participantes tienen poca despersonalización (43/105).
- Igual se interpretan el resto de frecuencias (Observamos que las frecuencias marginales por filas o columnas suman 1)

2. *Frecuencias condicionales.* En la tabla siguiente hemos calculado las frecuencias relativas condicionales respecto a cada fila y cada columna. Observamos que las frecuencias relativas respecto a cada fila o cada columna suman 1.

Frecuencias relativas por filas

Relaciones	Despersonalización		Total
	Bajo	Alto	
Muchas	0,52	0,48	1
Pocas	0,32	0,68	1

Frecuencias relativas por columnas

Relaciones	Despersonalización	
	Bajo	Alto
Muchas	0,56	0,35
Pocas	0,44	0,65
Total	1,00	1,00

Podemos ahora interpretar estas frecuencias.

- Porcentajes por filas
 - El 52% de personas *con muchas relaciones*, tienen baja despersonalización y el 48% de personas con muchas relaciones, tienen alta despersonalización.
 - El 32% de personas *con pocas relaciones*, tienen baja despersonalización y el 68% de personas con pocas relaciones, tienen alta despersonalización.
 - Por tanto es más probable tener alta despersonalización entre las que tiene pocas relaciones personales que entre las que tienen muchas. Los porcentajes en cada fila suman 100.
- Porcentajes por columna
 - El 56% de personas con *baja despersonalización*, tienen muchas relaciones y el 44% pocas relaciones
 - El 35% de personas con *alta despersonalización*, tienen muchas relaciones, mientras el 65% tienen pocas relaciones
 - Por tanto es más probable tener personas con pocas relaciones en las que tienen alta despersonalización. Los porcentajes de cada columna suman 100

Finalmente mostramos en una tabla todos los resultados anteriores

Relaciones personales	Despersonalización			Interpretación
	Bajo	Alto	TOTAL	
Muchas	24	22	46	Frecuencia absoluta
	0,23	0,21		Frecuencias relativas
	0,52	0,48		Respecto a fila
	0,56	0,35	0,4381	Respecto a columna
Pocas	19	40	59	Frecuencia absoluta
	0,18	0,38		Frecuencias relativas
	0,32	0,68		Respecto a fila
	0,44	0,65	0,5619	Respecto a columna
Total	43	62	105	
	0,4095	0,5905		Respecto a fila

Ejercicio 1.2. Idealismo

A continuación mostramos una tabla con el cruce idealismo, medida esta variable a partir del grado de acuerdo-desacuerdo de los entrevistados con la siguiente frase: “La vida sólo tiene sentido cuando una persona se dedica plenamente a una causa o ideal” y filiación política

Filiación política	La vida sólo tiene sentido por un ideal				Total	%
	muy desacuerdo	más bien desacuerdo	más bien desacuerdo	muy acuerdo		
Izquierda			420	73	954	58,1
Centro		106		36	298	
Derecha	20		181		389	23,7
Total	143	598		167	1641	
%	8,7	36,4	46,7	10,2	100	

Valores perdidos = 787

1. Rellena los huecos de la tabla, con las frecuencias y porcentajes que faltan
2. Interpreta los resultados obtenidos –presta atención a los 787 casos perdidos.

1. Completar la tabla

El método más eficaz para completar la tabla es buscar las filas y columnas que únicamente les falte un valor (incluido el valor del total).

- Primer paso. Completamos la fila del “Total” tenemos: que $143 + 598 + X + 167 = 1641$; por tanto $X = 1641 - 143 - 598 - 167 = 733$

Ideología recodificada	Vida sólo tiene sentido				Total	%
	muy desacuerdo	más bien desacuerdo	más bien desacuerdo	muy acuerdo		
Izquierda			420	73	954	58,1
Centro		106		36	298	
Derecha	20		181		389	23,7
Total	143	598	733	167	1641	
%	8,7	36,4	46,7	10,2	100	

- Segundo paso. Completamos las columnas “Muy acuerdo”; sabemos que $73 + 36 + X = 167$; siendo $X = 167 - 73 - 36 = 58$;
- Completamos la columna de “Más bien desacuerdo”; tenemos que $420 + X + 181 = 733$; siendo $X = 733 - 420 - 181 = 132$; La tabla queda:

Ideología recodificada	Vida sólo tiene sentido				Total	%
	muy desacuerdo	más bien desacuerdo	más bien desacuerdo	muy acuerdo		
Izquierda			420	73	954	58,1
Centro		106	132	36	298	
Derecha	20		181	58	389	23,7
Total	143	598	733	167	1641	
%	8,7	36,4	46,7	10,2	100	

- Tercer paso. Completamos la fila de “centro” tenemos: $X + 106 + 132 + 36 = 298$; por tanto, $X = 298 - 106 - 132 - 36 = 24$
- Completamos la fila de “derecha” tenemos: $20 + X + 181 + 58 = 389$,siendo $X = 389 - 20 - 181 - 58 = 130$. La tabla queda:

Ideología recodificada	Vida sólo tiene sentido				Total	%
	Muy desacuerdo	más bien desacuerdo	más bien desacuerdo	muy acuerdo		
Izquierda			420	73	954	58,1
Centro	24	106	132	36	298	
Derecha	20	130	181	58	389	23,7
Total	143	598	733	167	1641	
%	8,7	36,4	46,7	10,2	100	

- Cuarto paso, con el mismo procedimiento completamos la columna de “*Muy desacuerdo*” y la columna de “*Más bien desacuerdo*” La tabla queda:

Ideología recodificada	Vida sólo tiene sentido				Total	%
	muy desacuerdo	más bien desacuerdo	más bien desacuerdo	muy acuerdo		
Izquierda	99	362	420	73	954	58,1
Centro	24	106	132	36	298	
Derecha	20	130	181	58	389	23,7
Total	143	598	733	167	1641	
%	8,7	36,4	46,7	10,2	100	

- El último paso es completar la columna del “*Total*” tenemos en %: $58,1 + X + 23,7 = 100$ siendo $X = 100 - 58,1 - 23,7 = 18,2$. La tabla completa queda:

Ideología recodificada	Vida sólo tiene sentido				Total	%
	muy desacuerdo	más bien desacuerdo	más bien desacuerdo	muy acuerdo		
Izquierda	99	362	420	73	954	58,1
Centro	24	106	132	36	298	18,2
Derecha	20	130	181	58	389	23,7
Total	143	598	733	167	1641	
%	8,7	36,4	46,7	10,2	100	

Tema 2. Asociación estadística, dependencia funcional e independencia

Objetivos del tema 2

En este tema aprenderás a:

- Diferenciar la asociación estadística, dependencia funcional e independencia.
- Reconocer el tipo de relación entre dos variables de una tabla comparando las frecuencias condicionales
- Calcular las frecuencias esperadas en caso de independencia
- Analizar posibles explicaciones de una asociación estadística: relación causal, interdependencia, tercera variable explicativa o asociación espuria



Walter Frank Raphael Weldon

2.1. Introducción

En el tema anterior has aprendido a representar los datos en tablas de contingencia e interpretar las frecuencias absolutas dobles.

En muchas ocasiones al investigador le interesa saber si las dos variables presentadas en una tabla de contingencia están interrelacionadas entre sí. En este tema vamos a aclarar que entendemos por asociación estadística y a diferenciarla de la dependencia funcional y de la independencia. También aprenderemos a diferenciar entre relación causal y asociación y a estudiar diferentes situaciones que pueden llevar a la existencia de asociación.

2.2. Dependencia funcional y dependencia aleatoria

Comenzamos recordando la dependencia funcional que has estudiado en matemáticas y física. En algunos estudios donde se analizan dos variables observamos que, conociendo el valor de una de las variables podemos saber el valor exacto de la otra variable, ya que hay una fórmula matemática que las relaciona.

En la vida diaria podemos buscar ejemplos de estas dependencias funcionales, como es el caso de la relación entre espacio recorrido y velocidad de un móvil, para un intervalo de tiempo fijo. Como vemos en la tabla 2.1., dando valores distintos al espacio para un intervalo fijo de 5 horas, nos da como resultado un único valor para cada valor del espacio, de forma exacta, pues las variables están relacionadas por la ecuación:

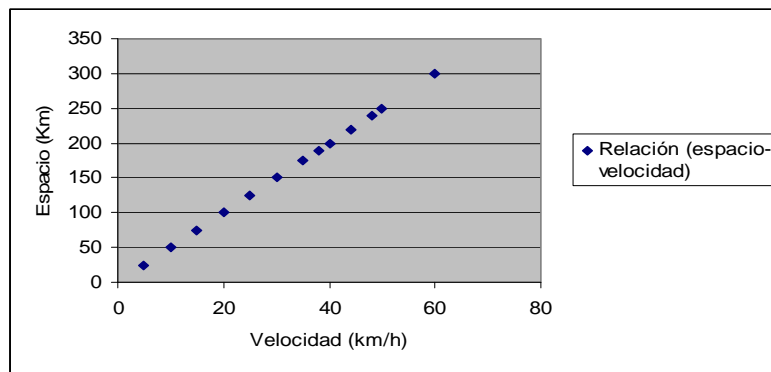
$$V = \frac{e}{t}; V = \text{velocidad}; e = \text{espacio}; t = \text{tiempo}$$

Tabla 2.1. Velocidad – espacio

Velocidad (Km/h)	5	10	15	20	25	30	35	38	40	44	48	50	60
Espacio (Km)	25	50	75	100	125	150	175	190	200	220	240	250	300

Observamos en la Figura 2.1. que en caso de dependencia funcional, al representar gráficamente las variables, obtenemos una línea que pasa por todos los puntos. En este tipo de relación, los valores que toma la variable Y quedan determinados, de un modo preciso, por los valores que toma la otra variable, que se considera como independiente.

Figura 2.1. Relación entre velocidad – espacio



2.3. Dependencia aleatoria

Existen casos en los que no se puede establecer con exactitud una fórmula que relacione dos variables, esto no quiere decir que no exista una cierta relación. Un caso es si queremos relacionar el coeficiente de inteligencia con las notas de una cierta prueba, es lógico pensar que cuanto más alto es el coeficiente de inteligencia la nota será mejor, aunque esto no ocurra en todos los casos o pueda variar ligeramente. Este tipo de dependencia llamaremos *dependencia aleatoria*. Si dichos pares de valores los representamos en un sistema cartesiano, los puntos, en general, no se ajustan de un modo preciso a una función matemática, sino que se obtiene un conjunto de puntos más o menos dispersos. Una representación de ese tipo recibe el nombre de *nube de puntos* o *diagrama de dispersión*.

El estudio de la posible relación entre dos variables cuantitativas suele iniciarse mediante la observación del correspondiente diagrama de dispersión o "nube de puntos". La presencia de una relación entre las variables se pondrá de manifiesto en el diagrama por una cierta tendencia de los puntos a acumularse en las proximidades de una línea. La relación será más intensa cuanto más concentrados estén los puntos y puede ser directa (si a mayor valor de una variable se da mayor valor de la otra) o inversa (a más valor de una variable se da menor valor de la otra).

2.4. El concepto de asociación

En algunos estudios nos interesa analizar si dos variables cualitativas están relacionadas entre sí. Para ello, se comienza estudiando los datos, representados en una tabla de contingencia. Al tratar de estudiar si existe o no una relación entre dos variables estadísticas, tratamos de contestar a las preguntas siguientes:

1. ¿Hay algún tipo de relación entre las variables?
2. ¿Podría medir la intensidad de esta relación mediante un coeficiente (coeficiente de asociación)?
3. ¿Sirve este coeficiente para poder comparar la intensidad de la relación de diferentes variables?
¿Cómo puedo interpretarlo?

En este tema estudiaremos la primera pregunta, y en los siguientes las preguntas 2 y 3.

Ejemplo 2.1. Trastorno digestivo

Se quiere estudiar si un cierto medicamento produce trastornos digestivos en los ancianos. Para ello se han observado durante un periodo de tiempo a 25 ancianos obteniendo los siguientes resultados.

<i>Sintomatología digestiva según se toma o no una medicina</i>			
	Molestias digestivas	No tiene molestias	Total
Toma la medicina	9	8	17
No la toma	7	1	8
Total	16	9	25

Utilizando los datos de la tabla, razona si en estos ancianos, el padecer trastornos digestivos está relacionado con haber tomado o no el medicamento, indica cómo has usado los datos.

Observa que, de los ancianos que toman la medicina, más o menos la mitad tienen molestias, pero de los que no la toman 7 de cada 8 las tienen. Por lo tanto, es más frecuente que tomen la medicina los que tienen molestias, o lo que es lo mismo las variables están asociadas. Si no lo estuvieran, lo que esperaríamos es que la proporción de los ancianos con molestias fuesen igual en los dos grupos.

En el ejemplo 2.1 vemos que, para concluir si las dos variables que forman una tabla de contingencia son dependientes o independientes se requiere un proceso de cálculo a partir de las frecuencias de la tabla. Podemos encontrar una variedad de situaciones, mostramos en la Figura 2.2. En el caso c) observamos que a cada valor de la variable X corresponde un solo valor de la variable Y y viceversa. Es claro que en este caso las variables están asociadas, y más aún se trataría de una *asociación perfecta* pues con toda seguridad podemos predecir el valor que tomara Y sabiendo el valor de X y viceversa.

Figura 2.2. Diferentes tipos de asociación

a) Independencia total					b) Asociación parcial					c) Asociación perfecta				
	Y ₁	Y ₂	Y ₃	Total		Y ₁	Y ₂	Y ₃	Total		Y ₁	Y ₂	Y ₃	Total
X ₁	10	20	70	100	X ₁	10	80	10	100	X ₁	100			100
X ₂	20	40	140	200	X ₂	80	20	100	200	X ₂			200	200
Total	30	60	210	300	Total	90	100	110	300	X ₃		100		100

En el caso a) *independencia*, observamos que las frecuencias absolutas dobles de los valores de Y_1, Y_2, Y_3 son proporcionales en X_1 y X_2 . Es decir, las frecuencias relativas condicionales de los valores de Y_1, Y_2, Y_3 son iguales en X_1 y X_2 . Diremos que las variables son independientes pues las frecuencias relativas condicionales de una de ellas no dependen del valor de la otra.

Los casos a) y c) tienen solo interés teórico pues lo más frecuente es encontrarse en el caso b) donde hay una *asociación parcial* en los datos, que sería el caso correspondiente al ejemplo. En este caso las frecuencias absolutas dobles de los valores de Y_1, Y_2, Y_3 no son proporcionales en X_1 y X_2 . Es decir, las frecuencias relativas condicionales de los valores de Y_1, Y_2, Y_3 son diferentes en X_1 y X_2 . Lo que nos va interesar desde el punto de vista estadístico es saber cuando esta relación sería estadísticamente significativa y cuál sería la intensidad de esta relación.

2.5. Independencia

Un método usual para conocer la independencia entre dos variables es mediante el estudio de las distribuciones condicionales. Para ello se calculan las distribuciones condicionales. En el caso de que coincidan, se puede afirmar que las dos variables son independientes.

Ejemplo 2.2. Insomnio y estrés.

Se pretende estudiar si el sufrir insomnio tiene relación con los trastornos de estrés. En una muestra de 250 personas observadas se obtuvieron los siguientes resultados:

	Padece estrés	No padece estrés	Total
Tener insomnio	90	60	150
No tener insomnio	60	40	100
Total	150	100	250

Aparentemente, las variables están relacionadas, pues la celda donde hay mayor frecuencia es en las personas que tienen estrés e insomnio. ¡Pero sería un error basar el *juicio de asociación* en los datos de una sola celda de la tabla!

Para analizar la asociación podemos comparar la proporción de personas con insomnio entre los que tienen estrés y los que no lo tienen. Obtendríamos la siguiente tabla de *frecuencias relativas condicionales por columnas*:

	Padece estrés	No padece estrés	Total
Tener insomnio	90/150=0,6	60/100=0,6	150/250=0,6
No tener insomnio	60/150=0,4	40/100=0,4	100/250=0,4
Total	150	100	250

En dicha tabla observamos que la proporción de personas con insomnio y sin insomnio es la misma entre los que tienen estrés y los que no lo tienen. Es decir *frecuencias relativas condicionales por columnas son iguales*.

Por otro lado, estas *frecuencias relativas condicionales por columnas* son también iguales a las *frecuencias relativas marginales por filas*, es decir a la proporción de personas con y sin insomnio en el total de la muestra.

En el ejemplo 2.2 podemos observar que la variable X es independiente de Y si todas las distribuciones de frecuencias relativas que se obtienen al condicionar X por diferentes valores de $Y = y_j$ son iguales entre sí e iguales a la distribución marginal de la variable X , es decir, cuando se verifica para todo par de valores i, j .

$$h(x_i/y_j) = h_i.$$

Esta propiedad significa que todas las distribuciones condicionales por columna coinciden con la distribución marginal de la variable X o lo que es lo mismo, la distribución de X no cambia cuando se condiciona por un valor de Y .

En el caso de independencia, se cumplen, además, las propiedades siguientes:

- La frecuencia relativa doble es igual al producto de las frecuencias relativas marginales de su fila y su columna: $h_{i,j} = h_i \cdot h_{.j}$, para todo i, j
- La frecuencia relativa marginal de cualquier valor de Y condicionada por un valor de X , es igual para todos los valores de X , es decir Y no depende de X : $h(y_j/x_i) = h_{.j}$. Podemos observarlo en el ejemplo calculando las frecuencias condicionales por filas, es decir, la proporción de personas con estrés y sin estrés dentro de las que tienen o no tienen insomnio:

	Padece estrés	No padece estrés	Total
Tener insomnio	90/150=0,6	60/150=0,4	150
No tener insomnio	60/100=0,6	40/100=0,4	100
Total	150/250=0,6	100/250=0,4	250

Vemos en el ejemplo que las proporciones de frecuencias con y sin estrés entre las que tienen o no insomnio son las mismas e igual que la proporción de personas con y sin insomnio en toda la muestra. Dicho de otro modo, las distribuciones condicionales por filas son todas iguales y coinciden con la distribución marginal de la variable Y , es decir que la distribución de Y no cambia cuando condicione por un valor de X .

2.6. Frecuencias esperadas en caso de independencia

La propiedad $h_{i,j} = h_i \times h_{.j}$, para todo i, j nos da un método de cálculo de las frecuencias teóricas en caso de independencia. Para ello, desarrollamos la fórmula anterior:

$$h_{i,j} = h_i \times h_{.j}; \text{ quiere decir que } \frac{f_{i,j}}{n} = \frac{f_{i.}}{n} \times \frac{f_{.j}}{n} \text{ y simplificando, obtenemos: } f_{i,j} = \frac{f_{i.} \times f_{.j}}{n}$$

Ejemplo 2.3. Insomnio y estrés (continuación)

Comprobaremos a continuación que en este ejemplo se dan las propiedades de independencia

- Si las variables son independientes, la frecuencia absoluta esperada en cada celda de la tabla es igual al producto de las frecuencias absolutas marginales de su fila y columna, dividido por el total de datos

Calculemos las frecuencias relativas dobles, comprobando que se cumple $h_{i,j} = h_{i.} \cdot h_{.j}$, para todo i, j

	Padece estrés (y_1)	No padece estrés (y_2)	Total
Tener insomnio (x_1)	0,36 (h_{11})	0,24 (h_{12})	0,6 ($h_{1.}$)
Tener insomnio (x_2)	0,24 (h_{21})	0,16 (h_{22})	0,4 ($h_{2.}$)
Total	0,6 ($h_{.1}$)	0,4 ($h_{.2}$)	1

Efectivamente, comprobamos que:

$$\begin{array}{llll}
 h_{11} = h_{1.} \cdot h_{.1} & 0,36 = 0,6 \times 0,6 & h_{12} = h_{1.} \cdot h_{.2} & 0,24 = 0,6 \times 0,4 \\
 h_{21} = h_{2.} \cdot h_{.1} & 0,24 = 0,4 \times 0,6 & h_{22} = h_{2.} \cdot h_{.2} & 0,16 = 0,4 \times 0,4
 \end{array}$$

- Calculemos las frecuencias esperadas en caso de independencia $e_{i,j} = \frac{f_{i.} \times f_{.j}}{n}$

	Padece estrés (y_1)	No padece estrés (y_2)	Total
Tener insomnio (x_1)	90 (f_{11})	60 (f_{12})	150 ($f_{1.}$)
Tener insomnio (x_2)	60 (f_{21})	40 (f_{22})	100 ($f_{2.}$)
Total	150 ($f_{.1}$)	100 ($f_{.2}$)	250 (n)

Aplicando la formula obtenemos que:

$$e_{1,1} = \frac{f_{1.} \times f_{.1}}{n} = \frac{150 \times 150}{250} = 90; \text{ que es precisamente la frecuencia observada en la celda}$$

$$f_{12} = f_{1.} \cdot f_{.2} / n \quad 60 = 150 \times 100 / 250$$

$$f_{21} = f_{2.} \cdot f_{.1} / n \quad 60 = 100 \times 150 / 250$$

$$f_{22} = f_{2.} \cdot f_{.2} / n \quad 40 = 100 \times 100 / 250$$

Como vemos en el ejemplo, las frecuencias observadas son iguales a las frecuencias esperadas en caso de independencia cuando las variables son independientes.

2.7. Tablas 2x2

En el caso de tablas con dos filas y dos columnas (tablas 2x2) podemos diferenciar entre dependencia directa y dependencia inversa. Veamos las celdas que nos informan del signo de la asociación en este caso.

Ejemplo 2.4. Alergia

Se quiere saber si sufrir o no de alergia tiene relación con llevar una vida sedentaria (llevar una vida sin realizar ningún tipo de ejercicio físico). Los datos de 300 sujetos se presentan a continuación.

Forma de vida	Sufre alergia	No sufre alergia
Sedentaria	130	30
No sedentaria	20	120

Observamos que hay 130 personas con alergia y vida sedentaria, es decir con los dos caracteres al mismo tiempo y asimismo, 120 personas sin alergia y con vida no sedentaria. En total 250 de las 300 personas de la muestra o tienen a la vez los dos caracteres o no tienen ninguno. Estas dos celdas (presencia-presencia y ausencia-ausencia) informan que la asociación en la tabla es directa.

Forma de vida	Sufre alergia	No sufre alergia
Sedentaria	Dep. directa	
No sedentaria		Dep. directa

Por el contrario en las otras dos celdas (presencia-ausencia y ausencia-presencia), se da un solo carácter y el otro no. Serían las celdas favorables a una asociación inversa. En el ejemplo hay solo 50 (30+20) casos en estas dos celdas y deducimos, en consecuencia que la asociación es directa.

Forma de vida	Sufre alergia	No sufre alergia
Sedentaria		Dep. inversa
No sedentaria	Dep. inversa	

2.7. Explicaciones de la asociación

El estudio de la asociación es un primer paso para describir relaciones de causalidad. Sin embargo, si encontramos asociación entre dos variables, ello no implica que una de ellas (A) sea causa de la segunda (B).

Ejemplo 2.5. Libros y coeficiente intelectual

Un estudio encontró una asociación positiva entre la cantidad de libros infantiles existentes en una casa y el coeficiente intelectual de los niños de la misma.

En aquellas casas donde abundan los libros infantiles el coeficiente intelectual de los niños es más alto que en aquellas casas donde hay muy pocos o ningún libro infantil.

Es muy común que al ver un estudio de este tipo, se infiera automáticamente que proveer de muchos libros infantiles a nuestros niños mejora su coeficiente intelectual.

Profundizando en este estudio, se encuentra que esa correlación positiva entre cantidad de libros e inteligencia, existe en forma independiente a si esos niños leen o no los libros.

La relación libros .vs. inteligencia es una correlación cierta, lo que no es cierta es la causalidad entre ellas. No es la cantidad de libros lo que hace a esos chicos más inteligentes, sino que hay un tercer factor, que hace por un lado que haya muchos libros en esa casa, y por otro que los niños de esa casa sean inteligentes.

Este tercer factor es el coeficiente intelectual de sus padres. Padres con alto coeficiente intelectual tienden a comprarles más libros a sus chicos. Padres con alto coeficiente intelectual tienden a tener hijos con alto coeficiente intelectual. No es entonces la cantidad de libros lo que hace inteligente a esos chicos, sino los genes que heredaron de sus padres.

Esto es un ejemplo muy ilustrativo de la común falacia lógica "correlación no implica causalidad"

Ejercicios resueltos del tema 2

Ejercicio 2.1. Decisión en un colegio

De los 650 padres del colegio A, 400 están en contra de que los niños vayan a colonias y el resto a favor, mientras que de los 260 padres del colegio B, 100 están a favor de que los niños realicen esta actividad y el resto en contra. Dispuestas las observaciones en una tabla de contingencia de 2 x 2 tenemos:

Muestras	Opinión		Total
	Favor	Contra	
Colegio A	250	400	650
Colegio B	100	160	260
Total	350	560	910

Estudia si hay asociación en estas variables.

Para analizar la asociación podemos comparar la proporción, como se vio durante el tema.

Obtendríamos la siguiente tabla de *frecuencias relativas condicionales por columnas*:

	Favor	Contra	Total
Colegio A	$250/350=0,714$	$400/560=0,714$	$650/910=0,714$
Colegio B	$100/350=0,286$	$160/560=0,286$	$260/910=0,286$
Total	1	1	1

En dicha tabla observamos que la proporción de personas del colegio A y del colegio B es la misma entre los que están a favor y los que están en contra. Es decir *frecuencias relativas condicionales por columnas son iguales*. Por tanto las variables son independientes.

Ejercicio 2.2. Caries en los niños

Para saber si la presentación de caries en niños está asociada con la experiencia de caries en al menos uno de los padres, se tomaron 523 niños de entre 12 y 15 años de edad y se les clasificó según su estado de caries dental (Baja, Normal y Alta) y según la experiencia de caries en sus padres (Baja, Normal y Alta), obteniéndose los datos de la tabla siguiente.

Niño	Padres		
	Baja	Normal	Alta
Baja	142	20	48
Normal	46	108	47
Alta	30	15	67

Estudia si hay asociación en estas variables.

Para analizar la asociación podemos comparar la proporción, como se vio durante el tema. Obtendríamos la siguiente tabla de *frecuencias relativas condicionales por columnas*:

Niño	Padres			Total
	Baja	Normal	Alta	
Baja	142/218= 0,651	20/143= 0,14	48/162= 0,3	210/523= 0,402
Normal	46/218= 0,211	108/143= 0,755	47/162= 0,29	201/523= 0,384
Alta	30/218= 0,138	15/143= 0,105	67/162= 0,41	112/523= 0,214
Total	1	1	1	1

Como las *frecuencias relativas condicionales por columnas son diferentes*, podemos concluir que hay asociación. Además, al ser en este caso las variables ordinales, podemos ver que la asociación es directa. Observamos que si los padres tienen baja experiencia en caries, la mayoría de los niños también tienen poca; cuando es normal, ocurre lo mismo e igualmente en alta. A mayor frecuencia de caries en los padres, mayor en los hijos. No podemos todavía medir la intensidad y tampoco decir la razón. Puede ocurrir que dependa de los hábitos higiénicos o bien de peor constitución de la dentadura, dietas pobres en calcio o una combinación de todos estos factores.

Ejercicio 2.3. Infarto de miocardio

Para evaluar si el ejercicio físico contribuye de algún modo a la prevención del infarto de miocardio, se tomaron 60 individuos que lo habían sufrido y otros 180 controles (sin infarto). También fueron clasificados como vigoroso si dedica más de 20 minutos/semana a algún ejercicio físico o no (si dedica menos). Los datos se presentan a continuación:

	No sufrido	Sufrido
no-vigorosos	25	75
vigorosos	35	105
Total	60	180

Estudia la asociación en las variables. Si son independientes comprueba las propiedades de independencia.

Para analizar la asociación podemos comparar la proporción, como se vio durante el tema. Obtendríamos la siguiente tabla de *frecuencias relativas condicionales por columnas*:

	No sufrido	Sufrido	Total
no-vigorosos	25/60=0,417	75/180=0,417	100/240=0,417
vigorosos	35/60=0,583	105/180=0,583	140/240=0,583
Total	1	1	1

Como las *frecuencias relativas condicionales por columnas son iguales*, podemos concluir con seguridad que son independientes. Se podría hacer de forma análoga con las frecuencias relativas por filas.

- Calculamos las frecuencias relativas dobles, comprobando que se cumple $h_{i,j} = h_{i.} \cdot h_{.j}$, para todo i, j

	No sufrido	Sufrido	Total
no-vigorosos	25/240=0,1042	75/240=0,3125	100/240=0,4167
Vigorosos	35/240=0,1458	105/240=0,4375	140/240=0,5833
Total	60/240=0,25	180/240=0,75	1

Efectivamente, comprobamos que:

$$\begin{aligned}
 h_{11} &= h_{1.} \cdot h_{.1} & 0,1042 &= 0,4167 \times 0,25 & h_{12} &= h_{1.} \cdot h_{.2} & 0,3125 &= 0,4167 \times 0,75 \\
 h_{21} &= h_{2.} \cdot h_{.1} & 0,1458 &= 0,5833 \times 0,25 & h_{22} &= h_{2.} \cdot h_{.2} & 0,4375 &= 0,5833 \times 0,75
 \end{aligned}$$

- Calculemos las frecuencias esperadas en caso de independencia $e_{i,j} = \frac{f_{i.} \times f_{.j}}{n}$

	No sufrido	Sufrido	Total
no-vigorous	25	75	100
Vigorous	35	105	140
Total	60	180	240

Aplicando la formula obtenemos que: $e_{1,1} = \frac{f_{1.} \times f_{.1}}{n} = \frac{100 \times 60}{240} = 25$; que es precisamente la frecuencia observada en la celda y lo mismo ocurre con los otros casos

$$f_{12} = f_{1.} \cdot f_{.2} / n \quad 75 = 100 \times 180 / 240$$

$$f_{21} = f_{2.} \cdot f_{.1} / n \quad 35 = 140 \times 60 / 240$$

$$f_{22} = f_{2.} \cdot f_{.2} / n \quad 105 = 140 \times 180 / 240$$

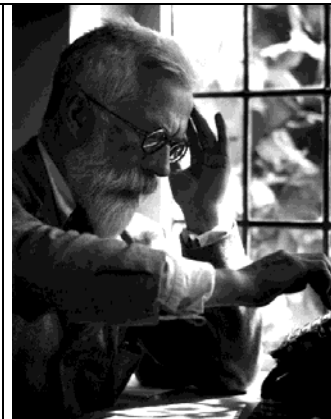
Las frecuencias observadas son iguales a las frecuencias esperadas en caso de independencia cuando las variables son independientes.

Tema 3. El estadístico Chi-cuadrado y contrastes asociados

Objetivos del tema 3

En este tema aprenderás a:

- Dar una medida de la diferencia entre frecuencias observadas y esperadas en caso de independencia en una tabla de contingencia
- Calcular e interpretar el estadístico Chi-cuadrado y sus grados de libertad
- Recordar los pasos y conceptos básicos en un contraste de hipótesis
- Comprender los pasos para llevar a cabo el contraste de independencia
- Contrastar la hipótesis de que dos variables en una tabla de contingencia son independientes, decidiendo si el valor Chi-cuadrado es estadísticamente significativo y tomando una decisión sobre el rechazo o de la hipótesis nula
- Interpretar el nivel de significación
- Realizar un contraste de Chi-cuadrado de homogeneidad para comprobar si varias subpoblaciones vienen de la misma población
- Comprender los supuestos de aplicación del contraste Chi-cuadrado



Ronald Aylmer Fisher

3.1. Introducción

Hasta ahora hemos analizado la existencia de asociación en los datos de una muestra dada, sin intentar extender las conclusiones a una población más amplia. En este tema estudiaremos la realización de una inferencia, donde se desea estudiar si la asociación encontrada entre dos variables en una muestra tomada al azar de una población mayor podría extenderse a la población de donde se tomaron los datos. Para ello, realizaremos un contraste de hipótesis.

Hay dos tipos de hipótesis que interesa contrastar, a partir de los datos de contingencia, el *contraste de homogeneidad* y el *contraste de independencia*. Los dos tipos de contrastes utilizan los datos de una tabla de contingencia y se basan en el estadístico Chi-cuadrado que estudiamos a continuación.

3.2. El estadístico Chi-cuadrado

Una medida muy extendida para medir la dependencia e independencia, es el estadístico Chi-cuadrado, que da una medida de la diferencia entre las frecuencias observadas en la tabla y las “frecuencias esperadas en caso de independencia”. Recordamos el cálculo de dichas frecuencias esperadas e_{ij} :

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

Con el estadístico Chi-cuadrado se obtiene una medida de diferencia entre las frecuencias esperadas y las frecuencias observadas. El estadístico se calcula en la forma siguiente:

$$\chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$$

Observamos las siguientes propiedades de este estadístico:

- Si todas las frecuencias observadas son iguales a la correspondiente frecuencia esperada, $f_{i,j} = e_{i,j}$

$$\text{entonces } \chi_{\text{exp}}^2 = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = \sum_i \sum_j \frac{(f_{ij} - f_{ij})^2}{e_{ij}} = 0.$$

- Esto ocurre sólo cuando las dos variables de la tabla son independientes; Por tanto, si hay independencia entre las dos variables de la tabla, $\chi_{\text{exp}}^2 = 0$

- Cuanto mayor sea la diferencia entre las frecuencias observadas y esperadas en la tabla, el valor de Chi cuadrado será mayor. Es decir, a mayor intensidad de la asociación entre las variables, Chi-cuadrado será mayor.
- El valor de Chi-cuadrado siempre es positivo o cero (pues es suma de números positivos, ya que los denominadores de la suma son todos positivos al ser suma de números elevados al cuadrado).
- En general, a mayor número de sumandos, se obtendrá un valor mayor.

Los *grados de libertad* de un estadístico calculado sobre un conjunto de datos se refieren al número de cantidades independientes que se necesitan en su cálculo, menos el número de restricciones que ligan a las observaciones y el estadístico. El número de grados de libertad del estadístico Chi-cuadrado se calcula de la siguiente forma:

- Se calcula, en primer lugar el número de sumandos, es decir $m \times n$, siendo n y m el número de filas y número de columnas en la tabla.
- A esta cantidad se debe restar el número de restricciones impuestas a las frecuencias observadas. Observamos que podemos cambiar todas las frecuencias de la tabla sin cambiar los totales por filas y columnas, excepto los datos en la última fila y la última columna de la tabla, pues una vez que fijemos todos los valores excepto estos, quedan automáticamente fijados. Por tanto, si la tabla tiene m filas y n columnas, el número de grados de libertad es $(m-1) \times (n-1)$. Expresamos esta dependencia en la siguiente forma:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \rightarrow \chi^2_{(n-1)(m-1)}$$

Ejemplo 3.1. Supervivencia en el Titanic

El 10 de abril de 1912, el Titanic zarpaba con 1317 pasajeros a bordo, ante la admiración de una muchedumbre de curiosos que contemplaban atónitos como aquella mole de acero se alejaba majestuosamente del puerto. Cinco días después los medios de comunicación de todo el mundo se hicieron eco de la increíble noticia: el barco más grande jamás construido yacía a casi cuatro mil metros de profundidad. La tabla 3.1 muestra la distribución de pasajeros, según supervivencia y tipo de pasaje

Tabla 3.1. Distribución de pasajeros en el Titanic según supervivencia y tipo de pasaje

	Sobrevive	No sobrevive	Total
Primera clase	194	128	322
Segunda clase	119	161	280
Tercera clase	138	573	711
Total	451	862	1313

Calculemos en el ejemplo las frecuencias esperadas en caso de independencia. Observamos que, una vez calculados los datos de la primera columna, los de la segunda se deducen automáticamente (es decir no son libres). Lo mismo ocurre con la última fila, una vez calculadas las dos primeras, queda automáticamente fijada. Por tanto los grados de libertad son $(3-1) \times (2-1) = 2 = k$ (denotamos como k . los grados de libertad)

$$e_{1,1} = \frac{f_{1.} \times f_{.1}}{n} = \frac{322 \times 451}{1313} = 110,6$$

En la tabla 3.2 mostramos las frecuencias esperadas en caso de independencia. Observamos que los grados de libertad son sólo 2, pues una vez calculadas una frecuencia esperada en la primera fila y otra en la segunda, las demás se deducen automáticamente, si no queremos variar los totales de filas y columnas.

Tabla 3.2. Frecuencias esperadas en el Titanic según supervivencia y clase social

	Sobrevive	No sobrevive	Total
Primera clase	110,6	211,4	322
Segunda clase	96,2	183,8	280
Tercera clase	244,2	466,8	711
Total	451	862	1313

Observamos, al comparar las tablas 3.1 y 3.2 que en primer clase hay mayor frecuencia observada que la esperada de supervivencia si no hubiese relación entre supervivencia y clase social. Mientras en segunda clase hay unos pocos más de lo esperado y en tercera casi la mitad de lo esperado. ¡El salvamento no fue entonces equitativo! A continuación llevamos a cabo los cálculos del estadístico Chi- cuadrado:

Tabla 3.3. Cálculo de Chi cuadrado

	Sobrevive	No sobrevive
Primera clase	$\frac{(194 - 110,6)^2}{110,6} = 62,9$	$\frac{(128 - 211,4)^2}{211,4} = 32,9$
Segunda clase	$\frac{(119 - 96,2)^2}{96,2} = 5,4$	$\frac{(181 - 183,8)^2}{183,8} = 2,8$
Tercera clase	$\frac{(138 - 244,2)^2}{244,2} = 46,2$	$\frac{(573 - 466,8)^2}{466,8} = 24,2$

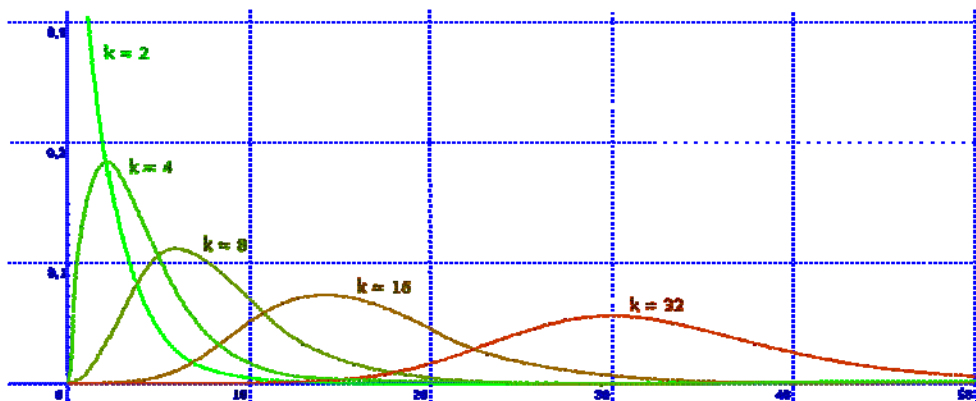
$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 62,9 + 32,9 + 5,4 + 2,8 + 46,2 + 24,2 = 174,4$$

Los grados de libertad, en este caso son $k = (3-1) \times (2-1) = 2$.

En la figura 3.1 mostramos la forma que toma el estadístico Chi-cuadrado, en caso de variables independientes, para diverso número de grados de libertad. Como hemos indicado, a mayor número de grados de libertad el valor será mayor. Así, para 4 grados de libertad la moda (valor más probable) se sitúa cerca del valor 5, mientras que para 32 grados de libertad se sitúa cerca de 39.

El valor obtenido 174,4 es muy poco probable en caso de independencia, pues observamos que para 2 grados de libertad los valores mayores que 10 apenas aparece. De hecho la probabilidad de obtener un valor mayor que 10,6 es sólo 0,005. Deducimos que el salvamento de los viajeros en el Titanic no fue independiente de su clase social.

Figura 3.1. Forma del estadístico Chi-cuadrado



Grados libertad	Probabilidad de un valor superior - Alfa (α)				
	0,1	0,05	0,025	0,01	0,005
1	2,71	3,84	5,02	6,63	7,88
2	4,61	5,99	7,38	9,21	10,60
3	6,25	7,81	9,35	11,34	12,84
4	7,78	9,49	11,14	13,28	14,86
5	9,24	11,07	12,83	15,09	16,75
6	10,64	12,59	14,45	16,81	18,55
7	12,02	14,07	16,01	18,48	20,28
8	13,36	15,51	17,53	20,09	21,95
9	14,68	16,92	19,02	21,67	23,59
10	15,99	18,31	20,48	23,21	25,19
11	17,28	19,68	21,92	24,73	26,76
12	18,55	21,03	23,34	26,22	28,30
13	19,81	22,36	24,74	27,69	29,82
14	21,06	23,68	26,12	29,14	31,32
15	22,31	25,00	27,49	30,58	32,80
16	23,54	26,30	28,85	32,00	34,27
17	24,77	27,59	30,19	33,41	35,72
18	25,99	28,87	31,53	34,81	37,16
19	27,20	30,14	32,85	36,19	38,58
20	28,41	31,41	34,17	37,57	40,00
21	29,62	32,67	35,48	38,93	41,40
22	30,81	33,92	36,78	40,29	42,80
23	32,01	35,17	38,08	41,64	44,18
24	33,20	36,42	39,36	42,98	45,56
25	34,38	37,65	40,65	44,31	46,93
26	35,56	38,89	41,92	45,64	48,29
27	36,74	40,11	43,19	46,96	49,65
28	37,92	41,34	44,46	48,28	50,99
29	39,09	42,56	45,72	49,59	52,34
30	40,26	43,77	46,98	50,89	53,67

3.3. Contraste de independencia

En el ejemplo 3.1 hemos llevado a cabo un contraste de independencia Chi-cuadrado, que nos permite determinar si existe una relación entre dos variables categóricas.

Recordarás que un *contraste de hipótesis* es un procedimiento estadístico, con una serie de pasos que lleva a la aceptación o rechazo de una hipótesis estadística. Los pasos a realizar en un contraste de hipótesis son los siguientes:

1. Fijar las hipótesis que se quieren contrastar: La hipótesis nula H_0 y la hipótesis alternativa H_1 . Estas hipótesis son complementarias una de otra.
2. Fijar el nivel de significación, o probabilidad máxima de rechazar la hipótesis nula H_0 , en caso de que sea cierta. Recordemos que el nivel de significación α es la probabilidad de Error Tipo I (probabilidad de rechazar la hipótesis nula, cuando de hecho es cierta).
3. Elegir un estadístico de contraste, que tenga alguna relación con la hipótesis. Formación a partir del estadístico de una regla de decisión, dividiendo los posibles valores del estadístico en dos regiones: (a) Si el estadístico cae en la región crítica (o de rechazo), se rechaza la hipótesis nula; (b) si el estadístico cae en la región de aceptación, no se puede rechazar la hipótesis nula.
4. Se comprueba el valor del estadístico y se toma la decisión de rechazar o no la hipótesis.

En el contraste de independencia, se desea decidir si las dos variables en una tabla de contingencia están o no asociadas. Siguiendo los pasos anteriores, se tendría

1. Fijar las hipótesis que se quieren contrastar. Estas hipótesis son las siguientes:

H_0 : Las variables en filas y columnas de la tabla son independientes

H_1 : Hay asociación entre las filas y columnas de la tabla

2. Fijamos el nivel de significación; lo más usual es elegir un valor $\alpha=0,05$. Esto quiere decir que la probabilidad máxima que fijamos para el error tipo I (rechazar la hipótesis de independencia cuando sea falsa) es 0,05.

3. Elegir un estadístico de contraste, que tenga alguna relación con la hipótesis. En este caso,

elegimos el estadístico Chi cuadrado, $\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \rightarrow \chi^2_{(n-1)(m-1)}$, que tiene relación

con la hipótesis nula, pues se basa en la comparación de frecuencias observadas y frecuencias esperadas en caso de independencia. Si la hipótesis nula H_0 es cierta (hay independencia entre filas y columnas) es de esperar un valor del Chi cuadrado será pequeño y si, por el contrario es falsa, será grande. Formaremos una regla decisión, dividiendo los posibles valores de Chi- cuadrado en dos regiones:

- Si el valor calculado χ^2_{exp} tiene una probabilidad menor que α (nivel de significación) rechazamos la hipótesis nula H_0 (hay independencia entre filas y columnas), pues el valor obtenido es improbable para una tabla con filas y columnas independientes. En este caso, suponemos que las variables están asociadas.
- Si el valor calculado χ^2_{exp} tiene una probabilidad igual o mayor que α (nivel de significación) no podemos rechazar la hipótesis nula H_0 . En este caso no tomamos ninguna decisión.

Nota: Observamos que el rechazo de la hipótesis nula tiene más fuerza que su aceptación, pues nos basamos en una situación muy poco probable: De ser cierta la independencia de las variables es muy poco probable obtener un alto valor de Chi- cuadrado. Por tanto, si obtenemos un alto valor de Chi- cuadrado, rechazamos que la hipótesis sea cierta.

Pero un valor pequeño de Chi cuadrado puede ser debido a varias causas: Puede ser que las variables sean independientes; puede ser que estén asociadas, pero la asociación sea muy pequeña; o puede ser que el tamaño de la muestra de datos sea pequeño y no permita ver la asociación. En este caso (cuando no podemos rechazar la hipótesis nula) tendríamos que estudiar mejor los datos para ver por qué se obtiene este valor pequeño de Chi- cuadrado.

Ejemplo 3.2. Deporte y bienestar

Un investigador quiere estudiar si hay asociación entre la práctica deportiva y la sensación de bienestar. Extrae una muestra aleatoria de 100 sujetos. Los datos aparecen a continuación.

Sensación de Bienestar	Práctica deportiva		Total
	Sí	no	
Sí	20	25	45
No	10	45	55
Total	30	70	100

Contraste la hipótesis de independencia entre bienestar y práctica de deporte (alfa = 0,01).

Primero calculamos las frecuencias esperadas en caso de independencia: $e_{ij} = \frac{f_i \cdot f_j}{n}$

Tabla 3.4. Frecuencias esperadas

Sensación de Bienestar	Práctica deportiva	
	Sí	No
Sí	13,5	31,5
No	16,5	38,5

Posteriormente calculamos el estadístico Chi-cuadrado:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 3,1296 + 2,5606 + 1,3413 + 1,0974 = 8,13$$

Los grados de libertad son: $(n-1) \times (m-1) = 1 \times 1 = 1$; Mirando en la tabla Chi-cuadrado obtenemos que la probabilidad de obtener un valor 8,13 o mayor con 1 grado de libertad es $p = 0,004$. Por tanto el valor es estadísticamente significativo, pues es menor que 0,01.

La decisión que se debe tomar es rechazar la hipótesis de independencia entre bienestar y práctica deportiva.

3.4. Contraste de homogeneidad

Otro caso en que usamos una tabla de contingencia es aquél en que se dispone de una población X clasificada en r subpoblaciones x_1, x_2, \dots, x_r . En cada una de estas poblaciones se toma una muestra, y los individuos de la misma se clasifican según una variable Y que puede tomar m valores posibles y_1, y_2, \dots, y_m . Sea p_{ij} la proporción de individuos que, en la población x_i tiene como valor de $Y=y_j$.

Un *contraste de homogeneidad* es cuando se desean contrastar las dos hipótesis siguientes:

- $H_0: p_{1j} = p_{2j} = \dots = p_{mj}$ para todo j ; dicho de otro modo, todas las subpoblaciones tienen idéntica distribución para la variable Y .
- H_1 : algunas de estas proporciones son diferentes. Dicho de otro modo, la distribución de la variable Y en alguna de estas subpoblaciones es diferente

El principal objetivo de realizar este contraste es comprobar que las distribuciones de todas las subpoblaciones son iguales o si hay alguna que difiere. Esto nos resulta práctico para poder combinar los resultados de todas las subpoblaciones, pues es necesario asegurarse de que los datos de las distintas muestras que se pretende agrupar son homogéneos.

Ejemplo 3.3. Grupo sanguíneo.

Se desea saber si la distribución de los grupos sanguíneos es similar en los individuos de dos poblaciones. Para ello se elige una muestra aleatoria de cada una de ellas, obteniéndose los siguientes datos ¿Qué decisión se debe tomar?

	A	B	AB	0	Total
Muestra 1	90	80	110	20	300
Muestra 2	200	180	240	30	650
Total	290	260	350	50	950

Calculamos las frecuencias esperadas:

Tabla 3.5. Frecuencias esperadas

	A	B	AB	0
Muestra 1	91.5789	82.105	110.53	15.789
Muestra 2	198.421	177.89	239.47	34.211

Posteriormente calculamos:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 1,76$$

Los grados de libertad son: $(n-1) \times (m-1) = 1 \times 3 = 3$

Mirando en la tabla Chi-cuadrado obtenemos que la probabilidad de obtener un valor 7,81 o mayor con 3 grado de libertad es $p = 0,184$. Por tanto el valor es no estadísticamente significativo, pues es mayor que 0,01. Aceptamos la hipótesis de homogeneidad de grupos sanguíneos en las dos muestras.

3.5. Interpretación y cálculo del p valor

El p-valor se puede interpretar de dos maneras diferentes:

- La probabilidad de rechazar la hipótesis nula cuando en verdad es cierta.
- La probabilidad de obtener un valor del estadístico igual o mayor al dado, cuando la hipótesis nula es cierta.

Esto significa en el caso de un contraste de independencia:

- Un valor cercano a $p=0$, indicaría un valor muy improbable de Chi-cuadrado si la hipótesis nula es cierta; por tanto llevaría a rechazar la hipótesis de independencia
- Un valor cercano a $p=1$, indicaría un valor muy probable de Chi-cuadrado si la hipótesis nula es cierta; por tanto no rechazaríamos la hipótesis de independencia

Cálculo del p valor:

Primero: los grados de libertad, $gl = (\text{filas}-1) \times (\text{columnas}-1)$.

Segundo: te sitúas en esos grados de libertad en la tabla (fila).

Tercero: buscas el valor de Chi- cuadrado de tu caso en la fila del segundo paso.

Cuarto: cuando lo sitúes, el valor de p será el que se indica en la parte superior de esa columna.

Por ejemplo, en el caso de grados de libertad = 1 y el valor del test sea 7,88, $p=0,005$.

Nota: Cuanto más alto es el valor de Chi cuadrado, más bajo es p-valor
--

3.6. Condiciones de aplicación de Chi- cuadrado

- Observa que al estudiar el valor de Chi-cuadrado en la tabla de la distribución, obtenemos siempre un valor positivo. Es decir, siempre hacemos un contraste unilateral.
- Si las frecuencias esperadas en las celdas son muy pequeñas, puesto que en la fórmula

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$$

aparecen dividiendo, se obtendría un valor alto de Chi-cuadrado, aunque las diferencias entre frecuencias observadas y esperadas fuese grande. Por eso, se recomienda que se use una muestra de suficiente tamaño. Estas son dos recomendaciones importantes

- Como máximo el 20% de las frecuencias esperadas pueden ser menores que el valor 5.
- No debe usarse si hay frecuencias esperadas inferiores a 1.

Ejercicios resueltos del tema 3

Ejercicio 3.1. Inmigración.

Se desea estudiar hasta qué punto existe relación entre el tiempo de residencia de inmigrantes en nuestro país y su percepción de integración. Se dispone de una muestra pequeña de 230 inmigrantes a los que se les evaluó en ambas variables obteniéndose la siguiente tabla de frecuencias observadas. ¿Confirman estos datos la hipótesis planteada con un nivel de confianza del 95%?

Tiempo de Residencia	Grado de integración		Total
	Bajo	Alto	
Más tiempo	40	90	130
Menos tiempo	90	10	100
Total	130	100	230

Calculamos las frecuencias esperadas: $e_{ij} = \frac{f_i \cdot f_j}{n}$, vemos que se cumplen las condiciones de aplicación de Chi-cuadrado.

Tabla 3.6. Frecuencias esperadas

Tiempo de residencia	Grado de integración	
	Bajo	Alto
Más tiempo	73,478	56,52
Menos tiempo	56,522	43,48



Posteriormente calculamos el estadístico Chi-cuadrado:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 80,69$$

Los grados de libertad son: $(n-1) \times (m-1) = 1 \times 1 = 1$. Al buscar en la tabla de Chi-cuadrado obtenemos un valor $p = 0,000$. Por tanto el Chi-cuadrado es estadísticamente significativo y rechazamos la hipótesis de independencia de las variables.

Tema 4. Medidas de asociación

Objetivos del tema 4

<p>En este tema aprenderás a:</p> <ul style="list-style-type: none"> • Interpretar la intensidad de la dependencia entre dos variables en una tabla de contingencia • Calcular e interpretar medidas de asociación en tablas 2x2: Coeficiente Phi de Pearson, Riesgo relativo y Razón de productos cruzados. • Calcular e interpretar medidas de asociación en tablas rxc: Coeficiente de contingencia de Pearson y V de Cramer. • Calcular e interpretar medidas de asociación que informan de la reducción del error de predicción de una variable, cuando se conoce el valor de la otra: Lambda de Goodman y Kruskal 	<p style="text-align: center;">Lambda de Goodman y Kruskal</p> <div style="display: flex; justify-content: space-around; align-items: center;">   </div> <p style="text-align: center;">1.- Leo A. Goodman</p> <p style="text-align: center;">2.- William Henry Kruskal</p>
---	--

4.1. Introducción

En el tema anterior hemos visto la manera de realizar un contraste Chi-cuadrado, para ver si dos variables están o no asociadas. Este contraste sólo nos informa de la presencia de asociación, pero no nos dice si es alta o baja.

En el caso de rechazar la hipótesis de independencia, el siguiente paso sería calcular algún valor que mida la *intensidad de la asociación*, es decir, indique cómo de dependientes son las variables la una de la otra. A continuación vamos a mostrar distintos coeficientes que sirven para medir esta intensidad. Diferenciamos el caso de las tablas 2x2 y los generales de tablas rxc.

4.2. Medidas para tablas 2x2

Las tablas 2x2 son especiales, porque, además de la intensidad de la asociación podemos ver el signo.

Recordamos la forma de estas tablas, que muestran la presencia o ausencia de dos factores *A* y *B*.

Tabla 4.1. Tabla de contingencia 2x2

	<i>B</i>	<i>No B</i>	<i>Total</i>
<i>A</i>	f_{11}	f_{12}	$f_{1\cdot}$
<i>No A</i>	f_{21}	f_{22}	$f_{2\cdot}$
<i>Total</i>	$f_{\cdot 1}$	$f_{\cdot 2}$	N

Para estas tablas podemos calcular algunos coeficientes con signo, de modo que se tiene:

- Si el *coeficiente es positivo la asociación es directa*, es decir, *A* y *B* suelen suceder juntos. Por tanto, si se da *A* suele darse *B*; por tanto habrá muchos casos en la celda f_{11} . Por otro lado, si no se da *A*, lo más frecuente es que tampoco se de *B*, por tanto habrá muchos casos en la celda f_{22} . Por ejemplo “ser rubio” y “ojos claros” tendría asociación directa, pues habrá muchos casos de rubios con ojos claros y también de morenos con ojos no claros.
- Si el coeficiente es negativo la asociación es inversa, es decir, si se da *A* no suele ocurrir *B* y si se da *B* no suele ocurrir *A*. Habría mayor frecuencia en las celdas f_{21} y f_{12} .
- Si el coeficiente es nulo no existe asociación, es decir, son independientes. No se encuentra un patrón en las diferentes celdas.

A continuación vamos a estudiar algunos coeficientes que pueden calcularse en una tabla 2x2, para saber no sólo el signo, sino la intensidad (si la asociación es fuerte o débil).

4.2.1. Coeficiente Phi de Pearson

Este coeficiente está basado en el valor chi-cuadrado, que vimos como se calculaba en el tema anterior. Recordamos que $\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$. Para las tablas 2x2 podemos utilizar la forma alternativa para el cálculo de Chi-cuadrado:

$$\chi^2_{\text{exp}} = \frac{(f_{11}f_{22} - f_{12}f_{21})^2 n}{f_{1\cdot} \cdot f_{2\cdot} \cdot f_{\cdot 1} \cdot f_{\cdot 2}}$$

Se define el coeficiente Phi, de la forma siguiente:

$$\Phi = \sqrt{\chi^2 / n} = \sqrt{\frac{((f_{11}f_{22} - f_{12}f_{21})^2 n) / (f_{1\cdot} \cdot f_{2\cdot} \cdot f_{\cdot 1} \cdot f_{\cdot 2})}{n}} = \sqrt{\frac{(f_{11}f_{22} - f_{12}f_{21})^2}{f_{1\cdot} \cdot f_{2\cdot} \cdot f_{\cdot 1} \cdot f_{\cdot 2}}}$$

Este coeficiente toma valores entre -1 y 1:

- El valor máximo (1) se obtiene cuando la dependencia es directa y perfecta, todos los casos están en las celdas f_{11} y f_{22} . Si el coeficiente es positivo, la dependencia es directa y más alta cuanto más se acerque a 1.
- El valor mínimo (-1) se obtiene cuando la dependencia es inversa y perfecta, todos los casos están en las celdas f_{12} y f_{21} . Si el coeficiente es negativo, la dependencia es inversa y más alta cuanto más se acerque a -1.
- El valor 0 se obtiene cuando hay independencia.
- Puede demostrarse que es equivalente al coeficiente de correlación cuando se codifican los valores A y B por 0 y no A y no B por 1
- No depende de las frecuencias marginales
- No varía si se multiplica o divide todas las frecuencias por el mismo número

Ejemplo 4.1. Observación de conducta

Para realizar un estudio de observación de conductas de interacción en niños en situación de juego se ha entrenado a dos observadores en la utilización de un sistema de registro de conductas. Los dos observadores codifican con el mismo sistema de categorías, requiriéndose que lo utilicen con un mismo criterio. Para evaluar el nivel de acuerdo entre los observadores y constatar si el entrenamiento recibido ha sido adecuado, se pide a ambos observadores que clasifiquen las conductas observadas en un vídeo de prueba. Los resultados fueron los siguientes:

Observador O1	Observador O2		Total
	A	B	
A	100	10	110
B	20	60	80
Total	120	70	190

Calcula e interpreta el coeficiente Phi de Pearson como valor de la intensidad.

Calculamos primero las frecuencias esperadas:

Tabla 4.2. Frecuencias esperadas

	A	B
A	$(110 \times 120) / 190 = 69,474$	$(110 \times 70) / 190 = 40,53$
B	$(80 \times 120) / 190 = 50,526$	$(80 \times 70) / 190 = 29,47$

A partir de ellas obtenemos el valor Chi-cuadrado:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = \frac{(100 - 69,474)^2}{69,474} + \frac{(20 - 50,526)^2}{50,526} + \frac{(10 - 40,53)^2}{40,53} + \frac{(60 - 29,47)^2}{29,47} = 13,413 + 18,443 + 22,99 + 31,62 = 86,47$$

Calculamos el coeficiente Phi de Pearson:

$$\Phi = \sqrt{\chi^2 / n} = \sqrt{86,47 / 190} = 0,675$$

Observamos que el valor es positivo (dependencia directa) y, en efecto, aparecen muchos más datos en la diagonal principal f_{11} y f_{22} que en la otra diagonal. Vemos que el valor es moderado-alto (cerca de 0,7).

Si tuviéramos un valor del observador O_1 , sabiendo el tipo de resultado, podríamos tratar de predecir su resultado por el observador O_2 . La mayoría de los que tienen un resultado A por el observador O_1 , también obtienen un resultado A por el observador O_2 (e igual con los resultados de B). Cuantos más casos en la diagonal principal, mayor valor de Phi y más fácil la predicción.

4.2.2. Riesgo relativo

Se puede calcular por filas y por columnas. El riesgo relativo

- El *riesgo relativo por columnas* indica cuanto más probable es la presencia de A con B que entre aquellos que no poseen B. Se calcula mediante la siguiente formula (columnas):

$$RR_{\text{columnas}} = \frac{P(A/B)}{P(A/\bar{B})} = \frac{f_{11} / f_{.1}}{f_{12} / f_{.2}} = \frac{f_{11} f_{.2}}{f_{.1} f_{12}}$$

- El *riesgo relativo por filas* indica cuanto más probable es la presencia de B con A que entre aquellos que no poseen A. Este coeficiente se calcula mediante la siguiente formula para filas:

$$RR_{\text{filas}} = \frac{P(B/A)}{P(B/\bar{A})} = \frac{f_{11} / f_{.1}}{f_{21} / f_{.2}} = \frac{f_{11} f_{.2}}{f_{21} f_{.1}}$$

Estos dos valores puede que coincidan, pero esto no pasa siempre.

El valor del Riesgo relativo cambia según el tipo de asociación que tengan las variables:

- El $RR = 1$, informa que no hay asociación entre las variables.
- El $RR > 1$, nos dice que existe asociación positiva.
- El $RR < 1$, indica que existe una asociación negativa.

Ejemplo 4.1. Observación de conducta (continuación)

Calculemos estos dos riesgos en el ejemplo sobre observación de conducta

Observador O_1	Observador O_2		Total
	A	B	
A	100	10	110
B	20	60	80
Total	120	70	190

El *riesgo relativo* por columnas se calcula mediante la siguiente formula:

$$RR_{columnas} = \frac{P(O_1 - A / O_2 - A)}{P(O_1 - A / O_2 - B)} = \frac{100/120}{10/70} = \frac{100 \times 70}{10 \times 120} = \frac{7000}{1200} = 5,8333$$

El $RR_{columnas} > 1$, nos dice que existe asociación positiva. Nos dice que es 5,8333 veces más fácil tener un valor A por el observador O_1 cuando se tiene un valor A por el observador O_2 que si se tiene un valor B por el observador O_2 .

El *riesgo relativo* por filas se calcula mediante la siguiente formula:

$$RR_{filas} = \frac{P(O_2 - A / O_1 - A)}{P(O_2 - A / O_1 - B)} = \frac{100/110}{20/80} = \frac{100 \times 80}{20 \times 110} = \frac{8000}{2200} = 3,6364$$

El $RR_{filas} > 1$, nos dice que existe asociación positiva. Nos dice que es 3,6364 veces más fácil tener un valor A por el observador O_2 cuando se tiene un valor A por el observador O_1 que si se tiene un valor B por el observador O_1 .

4.2.3. Razón de productos cruzados

Este coeficiente es una razón de dos cocientes:

$$RC = \frac{f_{11}f_{22}}{f_{12}f_{21}} = \frac{f_{11}/f_{21}}{f_{12}/f_{22}} = \frac{C_1}{C_2}$$

- C_1 es la razón de casos en que se presenta A y los que no se presenta A cuando está presente B.
- C_2 es la razón de casos A y no A cuando no está presente el factor B.

Conviene observar que la Razón de productos cruzados es una medida no simétrica. Es decir, A es la variable dependiente y B la independiente. Podemos interpretarlo en la forma siguiente:

- El $RC = 1$, implica que hay la misma razón de casos que aparece A y \bar{A} , cuando está B, que cuando no está presente B,
- El $RC < 1$, implica que la razón entre los casos que aparecen A y \bar{A} es menor cuando está presente B.
- El $RC > 1$, implica que la razón entre los casos que aparecen A y \bar{A} es mayor cuando está presente B.

Ejemplo 4.1. Observación de conducta (continuación)

Calculamos la razón de productos cruzado para este ejemplo que hemos estado utilizando

anteriormente, obteniendo: $RC = \frac{f_{11}f_{22}}{f_{12}f_{21}} = \frac{100 \times 60}{10 \times 20} = \frac{6000}{200} = 30$

Al obtener un valor mayor que 1, implica que la razón entre los resultados A y B del observador O_1 es superior cuando el sujeto tiene un valor A por el observador O_2 que cuando tiene un valor B.

De hecho, entre los sujetos con resultados A del observador O_2 hay 100 valorados con A por el observador O_1 por cada 20 valorados con B por el observador O_2 (la razón es 5/1). Entre los sujetos con resultados B del observador O_2 hay 10 valorados con A por el observador O_1 por cada 60 valorados con B por el observador O_2 .

4.3. Medidas de asociación para tablas rxc

Ahora vamos a mostrar algunos coeficientes los cuales se pueden aplicar a tablas con mayor número de columnas y filas.

4.3.1. Coeficiente de contingencia de Pearson

Este coeficiente se calcula mediante la siguiente formula:

$$C = \sqrt{\chi^2 / (\chi^2 + n)}$$

Vemos como se sigue basándose en el estadístico Chi-cuadrado. Este coeficiente cuando vale 0 indica independencia absoluta, pero el máximo, cuando la tabla tiene c columnas y r filas es:

$$\text{Max } \{C\} = \sqrt{\frac{\text{Min}\{r-1, c-1\}}{1 + \text{Min}\{r-1, c-1\}}}$$

Ejemplo 4.2. Aborto

Para analizar si el estado civil no era una variable relevante a la hora de explicar las actitudes abortistas, se ha encuestado a 500 sujetos obteniendo los resultados que aparecen en la tabla siguiente.

	Actitud Abortista	Actitud Antiabortista	Total
Solteros	120	30	150
Casados	50	200	250
Divorciados	30	70	100
Total	200	300	500

Calcula e interpreta el coeficiente de contingencia de Pearson y su máximo.

Calculamos las frecuencias esperadas y vemos que hay más solteros con actitud abortista de lo esperado (y menos contrarios al aborto); hay menos casados con esta actitud que lo esperado y algunos divorciados más. Por tanto hay relación entre las variables.

Tabla 4.3. Frecuencias esperadas

	Actitud Abortista	Actitud antiabortista
Solteros	$(150 \times 200) / 500 = 60$	$(150 \times 300) / 500 = 90$
Casados	$(250 \times 200) / 500 = 100$	$(250 \times 300) / 500 = 150$
Divorciados	$(100 \times 200) / 500 = 40$	$(100 \times 300) / 500 = 60$

Al calcular el valor Chi-cuadrado, se obtiene un valor estadísticamente significativo (se puede comprobar mirando la tabla para 2 grados de libertad:

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 60 + 40 + 25 + 16,67 + 2,5 + 1,667 = 145,83$$

Y el coeficiente C se calcula mediante la siguiente formula:

$$C = \sqrt{\chi^2 / (\chi^2 + n)} = \sqrt{145,83 / (145,83 + 500)} = 0,475$$

El máximo posible para esta tabla sería

$$\text{Max } \{C\} = \sqrt{\frac{\text{Min}\{r-1, c-1\}}{1 + \text{Min}\{r-1, c-1\}}} = 0,7071$$

Por ello, podemos decir que la asociación es moderada-alta.

4.3.2. V de Cramer

Este coeficiente se calcula mediante la siguiente formula:

$$V = \sqrt{\chi^2 / n(p-1)}$$

Siendo $p = \text{Min}\{\text{número de filas, número de columnas}\}$.

Este coeficiente varía entre 0 y 1, siendo 0 en caso de independencia y 1 en caso de dependencia perfecta.

Ejemplo 4.5. Aborto (continuación)

Para el ejemplo dado El valor chi-cuadrado era 145,83. El coeficiente se calcula mediante la siguiente formula, siendo $p = \text{Min}\{\text{número de filas, número de columnas}\}$:

$$V = \sqrt{\chi^2 / n(p-1)} = \sqrt{145,83 / 500 \times (2-1)} = \sqrt{0,29166} = 0,54$$

4.3.3. Medidas basadas en la reducción proporcional del error

Puesto que los coeficientes anteriores a veces no tienen una interpretación sencilla, algunos autores consideran medidas de asociación basadas en la cuantificación de la reducción del error que se comete al predecir el valor de una variable, cuando se conoce el valor de la otra. Supongamos que quiero predecir el valor de la característica X (variable en filas) en un individuo tomado al azar en la población.

- Si no tuviera ninguna información sobre el mismo, la probabilidad de cometer un error en la clasificación sería: $P(\varepsilon_1) = \frac{n - f_{\max}}{n}$, siendo f_{\max} la mayor frecuencia marginal en filas. Ya que si no se nada, lo asignaría a la fila de mayor frecuencia.
- Si supiera cuál es la columna (valor de la variable Y), lo asignaría a la fila (valor de X) que tenga frecuencia máxima en esa columna. La probabilidad de error sería $P(\varepsilon_2) = \frac{n - \sum f_{mj}}{n}$ donde f_{mj} es la frecuencia máxima en la columna j.
- El coeficiente $\lambda = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1}$ indica cual es el porcentaje de error que se ve reducido al predecir el valor de la variable dependiente X, conocido el valor de la variable independiente Y, en vez de asignar al azar el valor de X.

Lambda de Goodman y Kruskal

Una medida basada en la reducción proporcional del error es la lambda de Goodman y Kruskal, este coeficiente se calcula mediante la siguiente formula, para variables fila dependiente:

$$\lambda = \frac{\left(\frac{n - f_{\max}}{n}\right) - \left(\frac{n - \sum f_{mj}}{n}\right)}{\frac{n - f_{\max}}{n}} = \frac{(n - f_{\max}) - (n - \sum f_{mj})}{n - f_{\max}} = \frac{\sum f_{mj} - f_{\max}}{n - f_{\max}}$$

$$\text{Es decir, } \lambda = \frac{\sum f_{mj} - f_{\max}}{n - f_{\max}}$$

Siendo: f_{\max} es la mayor frecuencia marginal en filas y f_{mj} es la frecuencia máxima en la columna j-ésima.

Ejemplo 4.6. Aborto (continuación)

Recordamos los datos de la tabla.

	Actitud abortista	Actitud antiabortista	Total
Solteros	120	30	150
Casados	50	200	250
Divorciados	30	70	100
Total	200	300	500

- Supongamos que queremos predecir el estado civil de una persona al azar en esta muestra, sin saber nada sobre su actitud respecto al aborto. Como la máxima frecuencia marginal en filas es la correspondiente a casados (250), diríamos que es casado. La probabilidad de error en este caso sería $P(\varepsilon_1) = \frac{n - f_{\max}}{n} = \frac{500 - 250}{500} = 1/2$; pues nos equivocaríamos con los solteros y divorciados.
- Si nos dicen cuál es la actitud de la persona, predeciríamos que es soltero (si tiene actitud abortista) o casado (si no la tiene). La probabilidad de error ahora sería $P(\varepsilon_2) = \frac{n - \sum f_{mj}}{n} = \frac{500 - (120 + 200)}{500} = \frac{180}{500} = 0,36$. Hemos reducido el error de predicción, ahora sólo nos equivocamos en el 36% de los casos, en vez de la mitad
- El cociente $\lambda = \frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1} = \frac{0,5 - 0,36}{0,5} = \frac{0,14}{0,5} = 0,28$

En resumen, el coeficiente Lambda de Goodman y Kruskal se calcula mediante la siguiente formula:

$$\lambda = \frac{\sum f_{mj} - f_{\max}}{n - f_{\max}} = \frac{(120 + 200) - 250}{500 - 250} = 0,28$$

Siendo: f_{m+} es la mayor frecuencia marginal en filas y f_{mj} es la mayor frecuencia en la columna j-ésima. Hemos reducido un 28% con respecto al error que teníamos.

Ejercicios resueltos del tema 4

Ejercicio 4.1. Calificaciones de estadística.

Al final del curso los cuatro profesores de un mismo curso de estadística calificaron a los alumnos de la siguiente forma:

Profesores	NOTAS		Total
	Aprobados	Suspensos	
A	68	57	125
B	90	60	150
C	70	30	100
D	120	30	150
Total	348	177	525

- Calcula los coeficientes de contingencia de Pearson V de Cramer y Lambda de Goodman y Kruskal (con filas dependientes)
- ¿Qué se puede concluir de lo obtenido?

Calculamos primero las frecuencias esperadas:

Profesores	NOTAS	
	Aprobados	Suspensos
A	$(125 \times 348) / 525 = 82,86$	$(125 \times 177) / 525 = 42,14$
B	$(150 \times 348) / 525 = 99,43$	$(150 \times 177) / 525 = 50,57$
C	$(100 \times 348) / 525 = 66,29$	$(100 \times 177) / 525 = 33,71$
D	$(150 \times 348) / 525 = 99,43$	$(150 \times 177) / 525 = 50,57$

- El valor chi-cuadrado es: $\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 23,80$. Los grados de libertad son 3. Para 3 grados de libertad, la tabla me indica que el valor 12,84 tiene una probabilidad menor que 0,005. Por tanto el contraste es estadísticamente significativo y hay asociación entre profesor y aprobados/suspensos.
- El coeficiente de contingencia de Pearson sería $C = \sqrt{\chi^2 / (\chi^2 + n)} = 0,208$. El valor máximo en este caso es: $\text{Max } \{C\} = \sqrt{\frac{\text{Min}\{r-1, c-1\}}{1 + \text{Min}\{r-1, c-1\}}} = 0,7071$. Por tanto la intensidad de la asociación es pequeña pues sólo llegamos a la tercera parte del máximo.
- El coeficiente V de Cramer es: $V = \sqrt{\chi^2 / n(p-1)} = 0,213$. Como sabemos este coeficiente puede llegar a 1, por lo que confirma que la asociación es pequeña
- La lambda de Goodman y Kruskal: $\lambda_x = \frac{(\sum f_{mj}) - f_{m+}}{N - f_{m+}} = 0,08$, de lo que observamos que no mejora mucho la predicción al saber si el alumno ha aprobado para predecir su profesor.

Se puede concluir que todos los coeficientes están por debajo de la mitad del rango que pueden tomar, sin llegar a ser nula. Podría decirse que resulta una asociación baja.

Ejercicio 4.2. Niños disléxicos

Un psicólogo escolar está probando la eficacia de tres métodos de lectura de niños disléxicos. Estudiando los antecedentes de estos niños vio que la experiencia en la familia de hermanos disléxicos podría ser una causa de la ineficacia de los métodos. Con el fin de obtener alguna evidencia, midió a sus alumnos en las dos variables y obtuvo la tabla siguiente de curación.

	Pacientes curados			Total
	Método A	Método B	Método C	
Sin hermanos disléxicos	0	6	14	20
Con hermanos disléxicos	10	16	4	30
Total	10	22	18	50

- Calcula el coeficiente de contingencia de Pearson.
- Calcula el coeficiente V de Cramer.
- Calcula la lambda de Goodman y Kruskal.
- ¿Qué se puede concluir de lo obtenido?

Frecuencias esperadas:

	Método A	Método B	Método C
Sin hermanos disléxicos	$(20 \times 10) / 50 = 4$	$(20 \times 22) / 50 = 8,8$	$(20 \times 18) / 50 = 7,2$
Con hermanos disléxicos	$(30 \times 10) / 50 = 6$	$(30 \times 22) / 50 = 13,2$	$(30 \times 18) / 50 = 10,8$

- El valor chi-cuadrado es: $\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}} = 18,86$
- El coeficiente de contingencia de Pearson: $C = \sqrt{\chi^2 / (\chi^2 + n)} = 0,523$

El máximo es: $\text{Max}\{C\} = \sqrt{\frac{\text{Min}\{r-1, c-1\}}{1 + \text{Min}\{r-1, c-1\}}} = 0,7071$

- El coeficiente V de Cramer es: $V = \sqrt{\chi^2 / n(p-1)} = 0,614$
- La lambda de Goodman y Kruskal: $\lambda_x = \frac{(\sum f_{mj}) - f_{m+}}{N - f_{m+}} = 0,5$

Se puede concluir que todos los coeficientes están por encima de la mitad del rango que pueden tomar, sin llegar a ser el máximo. Podría decirse que resulta una asociación moderada-alta.



ANEXO 6
DESCRIPCIÓN DEL PROGRAMA DE
CÁLCULO EXCEL



El programa Excel permite la utilización de varias ventanas, que facilitan los cálculos y resultados. Hemos utilizado cinco divisiones: “Frecuencias”, “Gráficos”, “Test Chi cuadrado”, “Medidas asociación tablas 2x2” y “Medidas asociación tablas rxc”. Esta división se encuentra en la parte baja del programa cuando lo abrimos y analizaremos cada una en los puntos siguientes

1. Frecuencias

La primera pestaña es la que se llama “Frecuencias”, es la principal del programa. A partir de ella tendremos que introducir los nombres, valores y frecuencias dobles de las variables que queramos analizar. Todos los resultados de las demás pestañas dependerán de la correcta interpretación de lo que hagamos en esta. Una visión general de esta pestaña se presenta en la Figura 1.1:

Figura 1.1. Pestaña “Frecuencias”

FRECUENCIAS DOBLES								FRECUENCIAS RELATIVAS POR FILAS							
		Variable Y									Variable Y				
		Supervivencia									Supervivencia				
		y1	y2	y3	y4	y5			y1	y2	y3	y4	y5		
		Si	No	0	0	0	Total			Si	No	0	0	0	Total
Variable X	x1 Primera	194	128				322	Variable X	x1 Primera	0,60	0,40	0,00	0,00	0,00	1
Clase	x2 Segunda	119	161				280	Clase	x2 Segunda	0,43	0,58	0,00	0,00	0,00	1
	x3 Tercera	138	573				711		x3 Tercera	0,19	0,81	0,00	0,00	0,00	1
	x4						0		x4	0,00	0,00	0,00	0,00	0,00	0
	x5						0		x5	0,00	0,00	0,00	0,00	0,00	0
	Total	451	862	0	0	0	1313								
FRECUENCIAS RELATIVAS DOBLES								FRECUENCIAS RELATIVAS POR COLUMNAS							
		Variable Y									Variable Y				
		Supervivencia									Supervivencia				
		y1	y2	y3	y4	y5			y1	y2	y3	y4	y5		
		Si	No	0	0	0	Total			Si	No	0	0	0	Total
Variable X	x1 Primera	0,15	0,10	0,00	0,00	0,00	0,25	Variable X	x1 Primera	0,43	0,15	0,00	0,00	0,00	
Clase	x2 Segunda	0,09	0,12	0,00	0,00	0,00	0,21	Clase	x2 Segunda	0,26	0,19	0,00	0,00	0,00	
	x3 Tercera	0,11	0,44	0,00	0,00	0,00	0,54		x3 Tercera	0,31	0,66	0,00	0,00	0,00	
	x4	0	0,00	0,00	0,00	0,00	0,00		x4	0,00	0,00	0,00	0,00	0,00	
	x5	0	0,00	0,00	0,00	0,00	0,00		x5	0,00	0,00	0,00	0,00	0,00	
	Total	0,34	0,66	0,00	0,00	0,00	1,00	Total	1,00	1,00	0,00	0,00	0,00		

NOTA IMPORTANTE: SÓLO PUEDES ESCRIBIR EN LOS CUADROS VERDES

Algo importante a tener en cuenta es la nota marcada en amarillo, la cual nos informa de cuales son las celdas que tenemos que rellenar nosotros. Esto es válido en todas las pestañas del programa.

Datos requeridos:

Las celdas que tenemos que rellenar nosotros son las que tienen un fondo verde. Esto es válido en todas las pestañas del programa. En la Figura 1.1 podemos observar todas las celdas de datos que representan:

- Nombre de la variable X. El cual varía de un ejemplo a otro, por lo que requiere su introducción, en nuestro ejemplo es “Clase”.
- Nombre de la variable Y. El cual varía de un ejemplo a otro, por lo que requiere su introducción, en nuestro ejemplo es “Supervivencia”.
- Valores que pueden tomar las dos variables. El cual varía de un ejemplo a otro, por lo que requiere su introducción, en nuestro caso la variable X toma los valores ($x_1=Primera$, $x_2=Segunda$ y $x_3=Tercera$). La variable Y toma los valores ($y_1=Si$, $y_2=No$).
- Frecuencias absolutas dobles. El cual varía de un ejemplo a otro, por lo que requiere la

introducción de las frecuencias absolutas dobles, en nuestro ejemplo son 6 valores.

Resultados del programa:

- Frecuencias marginales. Las cuales aparecen en la Figura 1.2.
- Número total de sujetos en el experimento. Las cuales aparecen en la Figura 1.2.
- Tabla con las frecuencias relativas dobles, con sus totales. Las cuales aparecen en la Figura 1.3
- Tabla con las frecuencias relativas por columnas, con sus totales. Las cuales aparecen en la Figura 1.4.
- Tabla con las frecuencias relativas por filas, con sus totales. Las cuales aparecen en la Figura 1.5.

Objetivos:

- Resumir datos sobre las dos variables estadística que queremos analizar en una tabla de contingencia (Figura 1.2).
- Reconocer las dos variables estadísticas presentes en una tabla de contingencia. La variable X que representa las filas de la tabla, y la variable Y que representa las columnas de la tabla. En el ejemplo que aparece en la Figura 1.2., sobre supervivencia en el Titanic, la variable X es “Clase”, y la variable Y “Supervivencia”

Figura 1.2. Tabla de contingencia

FRECUENCIAS DOBLES							
		Variable Y					
		Supervivencia					
		y1	y2	y3	y4	y5	
		Si	No	0	0	0	Total
Variable X	x1 Primera	194	128				322
Clase	x2 Segunda	119	161				280
	x3 Tercera	138	573				711
	x4						0
	x5						0
	Total	451	862	0	0	0	1313

Figura 1.3. Frecuencia relativa doble

FRECUENCIAS RELATIVAS DOBLES							
		Variable Y					
		Supervivencia					
		y1	y2	y3	y4	y5	
		Si	No	0	0	0	Total
Variable X	x1 Primera	0,15	0,10	0,00	0,00	0,00	0,25
Clase	x2 Segunda	0,09	0,12	0,00	0,00	0,00	0,21
	x3 Tercera	0,11	0,44	0,00	0,00	0,00	0,54
	x4	0	0,00	0,00	0,00	0,00	0,00
	x5	0	0,00	0,00	0,00	0,00	0,00
	Total	0,34	0,66	0,00	0,00	0,00	1,00

Figura 1.4. Frecuencia condicional por columna

FRECUENCIAS RELATIVAS POR COLUMNAS							
		Variable Y					
		Supervivencia					
		y1	y2	y3	y4	y5	
		Si	No	0	0	0	
Variable X	x1 Primera	0,43	0,15	0,00	0,00	0,00	
Clase	x2 Segunda	0,26	0,19	0,00	0,00	0,00	
	x3 Tercera	0,31	0,66	0,00	0,00	0,00	
	x4	0	0,00	0,00	0,00	0,00	
	x5	0	0,00	0,00	0,00	0,00	
	Total	1,00	1,00	0,00	0,00	0,00	

Figura 1.5 Frecuencia condicional por fila

FRECUENCIAS RELATIVAS POR FILAS							
		Variable Y					
		Supervivencia					
		y1	y2	y3	y4	y5	
		Si	No	0	0	0	Total
Variable X	x1 Primera	0,60	0,40	0,00	0,00	0,00	1
Clase	x2 Segunda	0,43	0,58	0,00	0,00	0,00	1
	x3 Tercera	0,19	0,81	0,00	0,00	0,00	1
	x4	0	0,00	0,00	0,00	0,00	0
	x5	0	0,00	0,00	0,00	0,00	0

- Reconocer los valores que pueden tomar las dos variables estadísticas. La variable X puede tomar varios valores (x_1, x_2, \dots). La variable Y puede tomar varios valores (y_1, y_2, \dots). En el ejemplo que aparece en la Figura 1.2., la variable X toma los valores ($x_1=Primera, x_2=Segunda$ y $x_3=Tercera$). La variable Y toma los valores ($y_1=Si, y_2=No$).
- Identificar las frecuencias dobles que corresponde a cada par de valores (x_i, y_j) de las variables,

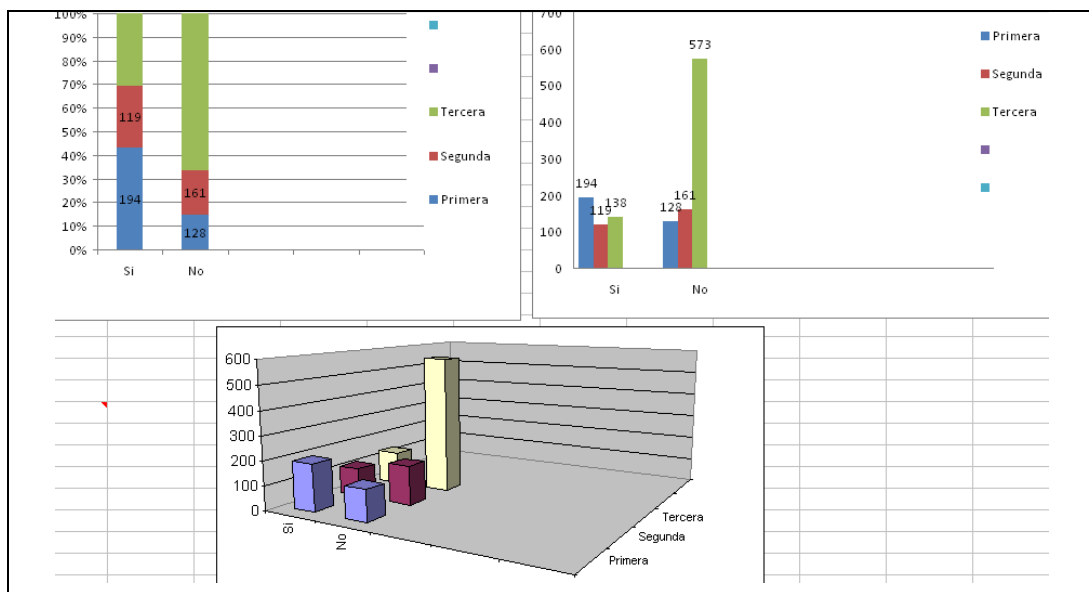
que representamos mediante f_{ij} . En el ejemplo que aparece en la Figura 1.2, podemos observar seis frecuencias absolutas: 194, 128, 119, 161, 138 y 573.

- Identificar el total de la muestra. En nuestro caso el valor es 1313. Calcular las frecuencias relativas dobles e interpretarlas. Si representamos mediante h_{ij} la frecuencia relativa del par de valores (x_i, y_j) , su valor se obtiene dividiendo cada celda por el total de pacientes n , es decir: $h_{ij}=f_{ij}/n$ (Figura 1.3). En el ejemplo obtenemos seis frecuencias relativas: 0,15, 0,1, 0,09, 0,12, 0,11 y 0,44.
- Obtener las frecuencias marginales e interpretarlas. Si en la tabla de frecuencias se suman las frecuencias por columnas, obtengo en cada columna j , el número de individuos $f_{.j}$ con un valor de la variable $Y=y_j$, independientemente del valor X . En el ejemplo podemos observar la frecuencia marginal de la variable X : 0,25, 0,21 y 0,54, y la frecuencia marginal de la variable Y : 0,34 y 0,66
- Obtener las frecuencias condicionales e interpretarlas. En caso de querer centrarnos en parte de los datos, podemos obtener $h(x_i/y_j)$ la frecuencia relativa condicional del valor x_i entre los individuos que presentan el carácter y_j (Figura 1.4.). Observamos que las frecuencias condicionales de cada columna suman 1. Podemos también obtener la frecuencia relativa de y_j condicionada por $x=x_i$ (Figura 1.5). Las frecuencias condicionales por filas también suman 1.

2. Gráficos

La segunda pestaña es la que se llama “Gráficos”, y nos aporta información visual. Está constituida por los gráficos relacionados con las tablas de contingencia. Una visión general de esta pestaña se presenta en la figura 2.1..

Figura 2.1. Pestaña “Gráficos”



Datos requeridos:

Como hemos dicho anteriormente, las celdas que tenemos que rellenar nosotros son las que tienen un fondo verde. Esto es válido en todas las pestañas del programa. En esta pestaña no hay ninguna de estas celdas, por lo que no se requieren datos.

Resultados del programa:

- Representación gráfica de los datos mediante diagrama de barras adosadas (Figura 2.2).
- Representación gráfica de los datos mediante diagrama de barras apiladas (Figura 2.3).
- Representación gráfica de los datos mediante gráfico tridimensional (Figura 2.4).

Objetivos:

- Observar e interpretar diagrama de barras adosadas. Este sistema de representación nos muestra con barras colocadas verticalmente la frecuencia, en porcentajes, de cada casilla del interior de la tabla de frecuencias (Figura 2.2.). Para cada valor de la variable X se presentan conjuntamente la frecuencia de cada valor de Y condicionado a este valor de X .
- Observar e interpretar los datos mediante diagrama de barras apiladas. Este gráfico muestra una barra por cada valor que tome la variable Y , las cuales a su vez, se dividen en distintos colores que representa a cada valor de la variable X . Representa la frecuencia con la que aparece cada valor de X en cada valor de Y , comparando entre categorías, el aporte de cada valor al total (Figura 2.3.).
- Observar e interpretar los datos mediante gráfico tridimensional. Este gráfico tridimensional coloca en el ancho y largo las variables X e Y , y la altura representa la frecuencia de cada celda (Figura 2.4).

Figura 2.2. Diagrama de barras adosadas

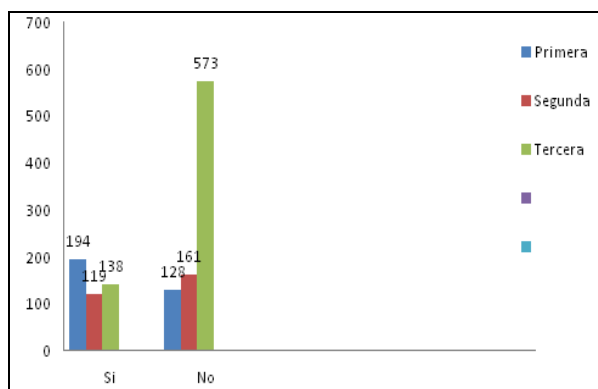


Figura 2.3. Diagrama de barras apiladas

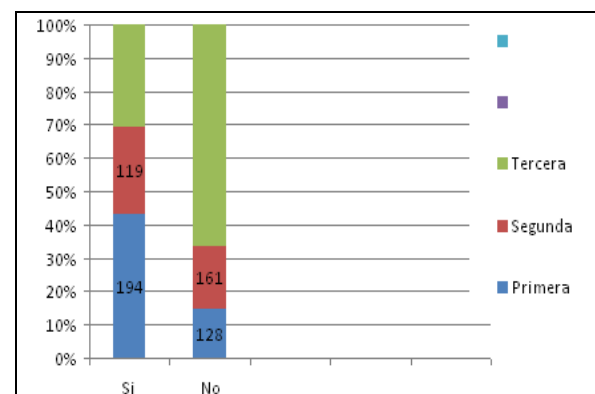
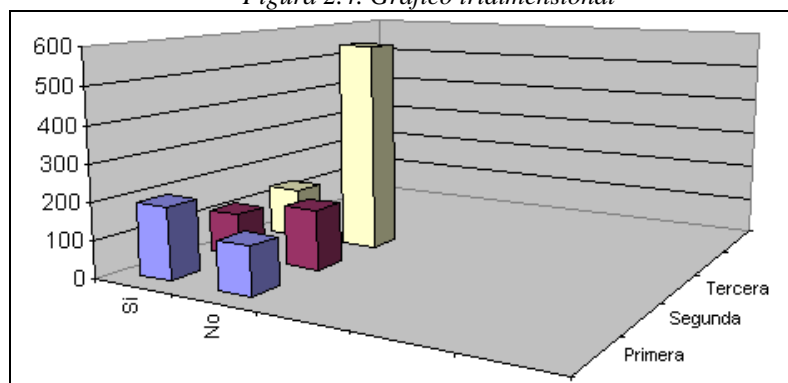


Figura 2.4. Gráfico tridimensional



3. Test Chi-cuadrado

La tercera pestaña es la que se llama “Test Chi cuadrado”, y es donde aparece este test tan importante en el estudio de estas tablas de contingencia. Una visión general de está pestaña se presenta en la figura 3.1.

Datos requeridos:

En este caso sólo hay una celda verde que tenemos que rellenar, que corresponde a los Grados de libertad.

Figura 3.1. Pestaña “Test Chi cuadrado”

FRECUCIAS OBSERVADAS								TEST CHI CUADRADO								
Variable Y								COMPONENTES DE CHI CUADRADO								
Supervivencia								Variable Y								
y1 y2 y3 y4 y5 Total								y1 y2 y3 y4 y5 Total								
Si No 0 0 0								Si No 0 0 0								
x1	Primera	194	128	0	0	0	322	Variable X	x1	Primera	62,883	32,9	0	0	0	95,783
x2	Segunda	119	161	0	0	0	280	Clase	x2	Segunda	5,4161	2,834	0	0	0	8,24982
x3	Tercera	138	573	0	0	0	711		x3	Tercera	46,199	24,17	0	0	0	70,3703
x4	0	0	0	0	0	0	0		x4	0	0	0	0	0	0	0
x5	0	0	0	0	0	0	0		x5	0	0	0	0	0	0	0
Total		451	862	0	0	0	1313		Total	114,5	59,91	0	0	0	174,40	
FRECUCIAS ESPERADAS EN CASO DE INDEPENDENCIA																
Variable Y																
Supervivencia																
y1 y2 y3 y4 y5 Total								Chi cuadrado= 174,40								
Si No 0 0 0								Grados liberta 2								
x1	Primera	110,6	211,4	0	0	0	322									
x2	Segunda	96,177	183,8	0	0	0	280									
x3	Tercera	244,22	466,8	0	0	0	711									
x4	0	0	0	0	0	0	0									
x5	0	0	0	0	0	0	0									
Total		451	862	0	0	0	1313									
										p=	0,000					

Resultados del programa:

- Tabla de frecuencias absolutas dobles, con las frecuencias marginales y el total de éstas (Figura 3.2).
- Tabla de frecuencias esperadas en caso de independencia (Figura 3.3).
- Tabla con las componentes de chi-cuadrado (Figura 3.4).
- Estadístico chi-cuadrado y el valor p (Figura 3.5).

Objetivos:

- Interpretar y reconocer el cálculo de las frecuencias esperadas en caso de independencia. La propiedad $h_{i,j} = h_{i.} \times h_{.j}$, para todo i, j nos da un método de cálculo de las frecuencias teóricas en caso de independencia. Desarrollamos esta fórmula, obtenemos: $f_{i,j} = \frac{f_{i.} \times f_{.j}}{n}$ (Figura 3.2). Cuando mayor sea la diferencia entre frecuencias observadas y esperadas mayor será la asociación.

Figura 3.2. Frecuencias esperadas en caso de independencia

FRECUENCIAS ESPERADAS EN CASO DE INDEPENDENCIA								
		Variable Y						
		Supervivencia						
		y1	y2	y3	y4	y5		
		Si	No	0	0	0	Total	
Variable X	x1	Primera	110,6	211,4	0	0	0	322
Clase	x2	Segunda	96,177	183,8	0	0	0	280
	x3	Tercera	244,22	466,8	0	0	0	711
	x4		0	0	0	0	0	0
	x5		0	0	0	0	0	0
		Total	451	862	0	0	0	1313

- Interpretar y reconocer el cálculo de las componentes de chi-cuadrado. Las componentes de Chi-cuadrado se obtienen mediante la formula $\frac{(f_{ij} - e_{ij})^2}{e_{ij}}$, las cuales se pueden representar en una tabla, la cual aparece en nuestro programa (Figura 3.3.). Las celdas que tengan mayor valor son las que más contribuyen a la asociación entre las variables.

Figura 3.3. Componentes de chi-cuadrado

TEST CHI CUADRADO								
COMPONENTES DE CHI CUADRADO								
		Variable Y						
		Supervivencia						
		y1	y2	y3	y4	y5		
		Si	No	0	0	0	Total	
Variable X	x1	Primera	62,883	32,9	0	0	0	95,783
Clase	x2	Segunda	5,4161	2,834	0	0	0	8,24982
	x3	Tercera	46,199	24,17	0	0	0	70,3703
	x4		0	0	0	0	0	0
	x5		0	0	0	0	0	0
		Total	114,5	59,91	0	0	0	174,40

- El calculo del estadístico chi-cuadrado. Con el estadístico Chi-cuadrado $\chi^2_{exp} = \sum_i \sum_j \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$ se obtiene una medida de diferencia entre las frecuencias esperadas y las frecuencias observadas (Figura 3.4).
- Cálculo de los grados de libertad. Los *grados de libertad* de un estadístico calculado sobre un conjunto datos se refieren al número de cantidades independientes que se necesitan en su cálculo, menos el número de restricciones que ligan a las observaciones y el estadístico. El número de grados de libertad del estadístico Chi-cuadrado es: $(m-1) \times (n-1)$.
- Interpretación del valor chi-cuadrado y del valor *p*. La interpretación dependerá de si realizamos un contraste de homogeneidad o de independencia.

Figura 3.4. Test Chi-cuadrado

Chi cuadrado=	174,40
Grados liberta:	2
p=	0,000

4. Medidas de asociación para tablas 2x2

La cuarta pestaña es la que se llama “Medidas asociación 2x2”. Esta pestaña sólo es utilizable en el caso de que nuestro ejercicio este constituido por una tabla de contingencia 2x2, y en ella encontraremos diferentes medidas de la intensidad de la asociación para este tipo de tablas. Una visión general se presenta en la figura 4.1.

Datos requeridos:

Como hemos dicho anteriormente, las celdas que tenemos que rellenar nosotros son las que tienen un fondo verde. Esto es válido en todas las pestañas del programa. En esta pestaña no hay ninguna de estas celdas, por lo que no se requieren datos.

Figura 4.1. Medidas de asociación 2x2

TABLA DE DATOS					
		Variable Y			
		Supervivencia			
		y1	y2		
		Si	No	Total	
Variable X	x1 Primera	194	128	322	
Clase	x2 Segunda	119	161	280	
	Total	313	289	602	
				Chi cuadrado	174,40
				Phi de Pearson	0,538
				Riesgo Relativo	1,39941094
					1,41761052
				Razón de productos cruzados	2,05055147

Resultados del programa:

- Tabla de contingencia 2x2, con sus totales.
- El valor chi-cuadrado como medida de asociación en una tabla 2x2.
- El valor Phi de Pearson como medida de asociación en una tabla 2x2.
- El valor riesgo relativo (por filas y columnas) como medida de asociación en una tabla 2x2.
- El valor razón de productos cruzados como medida de asociación en una tabla 2x2. Todos estos resultados se presentan en la figura 4.2.

Objetivos:

- Conocer el valor Phi de Pearson como medida de asociación en una tabla 2x2. Este coeficiente está basado en el valor chi-cuadrado, se define el coeficiente Phi de la forma siguiente: $\Phi = \sqrt{\chi^2 / n}$. Este coeficiente toma valores entre -1 (dependencia es inversa y perfecta) y 1 (dependencia es directa y perfecta).
- Conocer el valor riesgo relativo (por filas y columnas) como medida de asociación en una tabla 2x2. El *riesgo relativo por columnas* indica cuanto más probable es la presencia de A con B que entre aquellos que no poseen. El *riesgo relativo por filas* indica cuanto más probable es la presencia de B con A que entre aquellos que no poseen A.
- Conocer el valor razón de productos cruzados como medida de asociación en una tabla 2x2. Este coeficiente es una razón de dos cocientes: C1 es la razón de casos en que se presenta A y los que no se presenta A cuando está presente B, y C2 es la razón de casos A y no A cuando no está presente el factor B.

- Observar que el valor riesgo relativo por filas y por columnas no tiene que ser igual. Estos valores indican cosas diferentes, lo que significa que aunque pueden coincidir, no es lo que ocurre siempre.

Figura 4.2. Medidas de asociación 2x2

Chi cuadrado	174,40	
Phi de Pearson	0,538	
Riesgo Relativo	1,39941094	Columnas
	1,41761052	Filas
Razón de productos cruzados	2,05055147	

5. Medidas de asociación para tablas rxc

La quinta, y última pestaña es la que se llama “Medidas asociación rxc”. Esta pestaña sólo es utilizable en el caso de que nuestro ejercicio este constituido por una tabla de contingencia rxc, donde al menos una de estas medidas sea mayor de 2. En esta pestaña encontraremos diferentes medidas de la intensidad de la asociación para este tipo de tablas. Una visión general de está pestaña sería:

Figura 5.1. Medidas de asociación rxc

FRECUENCIAS DOBLES									
		Variable Y							
		Supervivencia							
		y1	y2	y3	y4	y5			
		Si	No	0	0	0	Total		
Variable X	x1 Primera	194	128	0	0	0	0	0	322
Clase	x2 Segunda	119	161	0	0	0	0	0	280
	x3 Tercera	138	573	0	0	0	0	0	711
	x4 0	0	0	0	0	0	0	0	0
	x5 0	0	0	0	0	0	0	0	0
	Total	451	862	0	0	0	0	0	1313
Número de filas		3							
Número de columnas		2							
Coeficiente de Contingencia de Pea				0,342	Máximo=	0,70710678			
Lambda de Goodman y Kruskal				0,093					
V de Cramer				0,364					

Datos requeridos:

Como hemos dicho anteriormente, las celdas que tenemos que rellenar nosotros son las que tienen un fondo verde. Esto es válido en todas las pestañas del programa.

- El número de filas y columnas que tiene la tabla (Figura 5.2.). Este valor es necesario para las formulas que se utilizan en esta pestaña, referidas al calculo de medidas de la asociación en tabla rxc.

Figura 5.2. Introducción del número de filas y columnas

Número de filas	3
Número de columnas	2

Resultados del programa:

- El valor coeficiente de contingencia de Pearson como medida de asociación en una tabla rxc. El valor Lambda de Goodman y Kruskal como medida de asociación en una tabla rxc.
- El valor V de Cramer como medida de asociación en una tabla rxc.
- El valor máximo del coeficiente de contingencia de Pearson. Todos ellos se presentan en la figura 5.3.

Objetivos:

- Calcular e interpretar el valor coeficiente de contingencia de Pearson como medida de asociación en una tabla rxc. Se sigue basando en el estadístico Chi-cuadrado. Este coeficiente cuando vale 0 indica independencia absoluta.

Figura 5.3. Medidas de asociación rxc

Coefficiente de Contingencia de Pea	0,342	Máximo=	0,70710678
Lambda de Goodman y Kruskal	0,093		
V de Cramer	0,364		

- Calcular e interpretar el valor Lambda de Goodman y Kruskal como medida de asociación en una tabla rxc. Es una medida basada en la reducción proporcional del error. A mayor valor indica que se reduce más el error cuando se predice una variable a partir de la otra.
- Calcular e interpretar el valor V de Cramer como medida de asociación en una tabla rxc. Este coeficiente se calcula mediante la siguiente formula: $V = \sqrt{\chi^2 / n(p-1)}$. Este coeficiente varía entre 0 y 1, siendo 0 en caso de independencia y 1 en caso de dependencia perfecta.
- Reconocer que el valor del coeficiente de contingencia de Pearson no es fijo. Este coeficiente cuando vale 0 indica independencia absoluta, pero el máximo, cuando la tabla tiene c columnas y r filas es: $Max \{C\} = \frac{Min\{r-1, c-1\}}{\sqrt{1+Min\{r-1, c-1\}}}$