



e-ISSN: 2595-0967

**Juicios de asociación en tablas de contingencia 2x2 por estudiantes de
Educación Media en Chile**

**Association judgments in contingency tables by High School students in
Chile**

**Julgamentos da associação em tabelas de contingência por estudantes do
Ensino Médio no Chile**

Nicolás Fernández Coronado
Universidad de Los Lagos
Osorno, Chile

nicolasalonso.fernandez@alumnos.ulagos.cl

Orcid: 0000-0002-9613-3144

Jaime I. García-García
Universidad de Los Lagos
Osorno, Chile

jaime.garcia@ulagos.cl

Orcid: 0000-0002-8799-5981

Daniela Calderón Torres
Universidad de Los Lagos
Osorno, Chile

danieladelpilar.calderon2@alumnos.ulagos.cl

Orcid:0000-0001-7036-3454

Elizabeth Hernández Arredondo
Universidad de Los Lagos
Osorno, Chile

elizabeth.hernandez@ulagos.cl

Orcid:0000-0002-5285-1603

Enviado:15/12/2020

Aceito:24/03/2021

DOI: 10.30612/tangram.v4i1.13176



Resumen: En este estudio exploratorio se analizan las respuestas de 75 estudiantes chilenos de Educación Media a la tarea de realizar juicios de asociación en tablas de contingencia 2x2 con variables cualitativas. Esto con el objetivo de identificar las concepciones iniciales que tienen del concepto de asociación, las estrategias que utilizan para justificar su respuesta y los errores que presentan en estas. Las justificaciones que dieron los estudiantes se clasificaron según su formalidad y con base en las categorías mencionadas por Estepa y Batanero (1995); además, se identificaron los conflictos semióticos señalados en un estudio previo por Cañadas, Batanero, Estepa y Arteaga (2013a). Como resultado, se identifican pocos estudiantes que logran responder a la tarea haciendo uso de las frecuencias o probabilidades, la presencia de conflictos semióticos como la concepción causal y la concepción determinista, y una alta cantidad de estudiantes de primeros años que no logra responder a la tarea.

Palabras clave: Tablas de doble entrada. Asociación. Educación Media.

Abstract: In this exploratory study, the answers of 75 Chilean high school students to the task of making association judgments in 2x2 contingency tables with qualitative variables are analyzed. The objective is to identify the initial conceptions they have of the concept of association, the strategies they use to justify their answers and the errors they present in these. The justifications given by the students were classified according to their formality and based on the categories mentioned by Estepa and Batanero (1995); in addition, the semiotic conflicts pointed out in a previous study by Cañadas, Batanero, Estepa and Arteaga (2013a) were identified. As a result, it is identified that few students achieve response to the task using frequencies or probabilities, the presence of semiotic conflicts such as causal conception and deterministic conception, and a high number of first year students who fail to respond to the task.

Keywords: Double entry tables. Association. Secondary Education.

Resumo: Este estudo exploratório analisa as respostas de 75 estudantes chilenos do ensino médio à tarefa de fazer julgamentos de associação em tabelas de contingência 2x2 com variáveis qualitativas. O objetivo é identificar as concepções iniciais que eles têm do conceito de associação, as estratégias que eles usam para justificar suas respostas e os erros que eles apresentam em suas respostas. As justificativas dadas pelos estudantes foram classificadas de acordo com sua formalidade e baseadas nas categorias mencionadas por Estepa e Batanero (1995); além disso, foram identificados os conflitos semióticos apontados em um estudo anterior de Cañadas, Batanero, Estepa e Arteaga (2013a). Como resultado,

identifica-se que poucos alunos conseguem responder à tarefa usando frequências ou probabilidades, a presença de conflitos semióticos como a concepção causal e a concepção determinista, e um alto número de alunos do primeiro ano que não conseguem responder à tarefa.

Palavras-chave: Tabelas de entrada dupla. Associação. Ensino Secundário.

Introducción

En las últimas décadas, la Estadística y Probabilidad han tomado un lugar importante dentro del currículo y los programas de estudio a nivel mundial, debido a su reconocimiento como herramientas claves en disciplinas científicas que las han vuelto partes fundamentales dentro de la investigación de carácter científico (Esquivel, 2016). La Estadística nos permite recolectar y manejar datos, mientras que la Probabilidad los sustentos para realizar inferencias con base en estos; es decir, ambas permiten el análisis y modelación de la realidad, generando así una interpretación sobre esta. Por ello, es necesario que el ciudadano tenga las capacidades para realizar dichas acciones y poder aplicarlas de manera efectiva dentro de su contexto, hoy más que nunca, debido a la situación de emergencia sanitaria por el coronavirus (SARS-CoV-2) que nos mantiene en diferentes estados de confinamiento y donde la cantidad de información que se recibe día a día de tipo estadística-probabilística generada por los medios de comunicación ha aumentado considerablemente, por lo que se ha vuelto necesario promover y priorizar los saberes que se consideran esenciales para los estudiantes (Vásquez, Ruz y Martínez, 2020).

Una de las formas más habituales en las que se presenta información del tipo estadística es la tabla de contingencia o de doble entrada, útil para la comunicación, representación y análisis de información, y que facilita la búsqueda de una correlación entre los datos que pueden representar fenómenos, características o realidades, expresada mediante lo conocido como un juicio de asociación (Cañadas, Contreras, Arteaga y Gea, 2013b). Sin embargo, a

pesar de que la realización de juicios de asociación en las tablas de contingencia es una práctica que abordan los aspectos priorizados antes mencionados, diversas investigaciones identifican la existencia de dificultades para su aprendizaje (e.g. Cañadas et al., 2013b; Cañadas, Arteaga, Guirado y Molina, 2017) y graves conflictos en las concepciones previas de los estudiantes (e.g. Estepa y Batanero, 1995; Gea, Batanero, López-Martín y Arteaga, 2016). El estudio de la correlación y juicios de asociación en estudiantes de secundaria ha sido abordado con un enfoque en la didáctica (Lavalle, Micheli y Rubio, 2006) y la experimentación (Estepa, 2007); sin embargo, no se ha profundizado en las concepciones iniciales y estrategias que estos usan al enfrentarse a este tipo de tareas, sin previa instrucción. Por estas razones, se sitúan los juicios previos de asociación de estudiantes de Educación Media, con base en tablas de contingencia, como nuestro objeto de estudio.

El objetivo del presente trabajo es explorar la competencia de estudiantes de Educación Media en Chile para la realización de juicios de asociación entre dos variables cualitativas representadas en una tabla de contingencia, sin previa instrucción, es decir, mediante sus concepciones iniciales; con el propósito de responder las siguientes interrogantes: ¿Qué estrategias utilizan los estudiantes chilenos de Educación Media para dar respuesta a la tarea de realizar juicios de asociación entre dos variables cualitativas representadas en tablas de contingencia? ¿Qué conflictos presentan los estudiantes al realizar dichos juicios de asociación? Para esto, se analizan las respuestas de 75 estudiantes de Educación Media a tres tareas sobre juicios de asociación en tablas de contingencia 2x2 sin previa instrucción, identificando las estrategias señaladas por Estepa y Batanero (1995), los conflictos indicados por Cañadas et al. (2013b) y Batanero, Godino y Estepa (1994). Esto permitirá reconocer aspectos esenciales para la elaboración de secuencias didácticas para la comprensión de la asociación de dos variables cualitativas en tablas de contingencia 2x2.

Fundamentos del estudio

En seguida, se presentan los fundamentos que sustentan esta investigación. Para ello, se muestran algunos resultados emergentes de las investigaciones acerca de este tipo de

problemas desde las siguientes dimensiones: 1) tablas de contingencia 2x2 y juicios de asociación, y 2) estrategias de resolución, concepciones previas sobre asociación estadística y conflictos semióticos. Apoyados en estos constructos teóricos nacidos de resultados empíricos, nos acercamos a mirar lo resultados de nuestro acercamiento.

Tablas de contingencia 2x2 y juicios de asociación

La tabla de contingencia es un instrumento de gran utilidad en la presentación y análisis de datos cualitativos, se utiliza con frecuencia en actividades profesionales, donde se estudia si existe asociación entre variables (Cañadas et al. 2017). Cañadas, Contreras, Gea y Arteaga (2012) introduce este tipo de representación como un componente importante de la cultura estadística que, dentro de su complejidad semiótica, posee interés en su aplicación en diversas disciplinas como la psicología.

El caso más simple corresponde a la tabla de dimensiones 2x2 (dos filas y dos columnas de frecuencias), que suele usarse para analizar datos estadísticos con respecto a dos variables (e.g. presentar síntomas y padecer la patología) que pueden tener valor verdadero o falso, distribuyéndose las frecuencias como se presenta en la Tabla 1.

Tabla 1: Modelo de tabla de contingencia 2x2

	A	~ A	TOTAL
B	a	b	$a + b$
~ B	c	d	$c + d$
TOTAL	$a + c$	$b + d$	$a + b + c + d$

Fuente: elaborada por los autores.

Además, Cañadas et al. (2012) indica que se pueden deducir y obtener los siguientes valores:

- Frecuencia relativa doble (probabilidad de que ocurran simultáneamente A y B):

$$\frac{a}{a+b+c+d}$$

- Frecuencia relativa condicional respecto a su fila (probabilidad condicional de A, dado que ocurrió B): $\frac{a}{a+b}$
- Frecuencia relativa condicional respecto a su columna (probabilidad condicional de B, dado que ocurrió A): $\frac{a}{a+c}$
- Frecuencia relativa marginal por fila (probabilidad simple de B): $\frac{a+b}{a+b+c+d}$
- Frecuencia relativa marginal por columna (probabilidad simple de A): $\frac{a+c}{a+b+c+d}$

Uno de los casos particulares de la Estadística es realizar un juicio de asociación entre dos variables. En tablas de contingencia 2x2 es necesario realizar operaciones y establecer relaciones entre las distintas frecuencias contenidas en la tabla, con tal de identificar si dos variables están o no relacionadas y, en caso de que lo estén, señalar de qué manera (fuerza de la correlación). Esto la vuelve una de las actividades cognitivas de mayor importancia respecto a fenómenos bivariados, al permitir explicar fenómenos del pasado, controlar variables en experimentos e investigaciones y predecir tendencias de sucesos futuros (Batanero, Cañadas, Contreras y Gea, 2015).

Respecto a esta práctica, Batanero et al. (1998) identificaron conflictos previos que se repetían con regularidad en estudiantes de nivel universitario (e.g. considerar una relación causal como la existencia de dependencia), algunos de los cuales no se lograron resolver con la intervención del profesor y la discusión con el alumno, los cuales pueden mantenerse en las estrategias resolutivas empleadas en la adultez (Cañadas et al., 2013b).

Estrategias de resolución, concepciones previas sobre asociación estadística y conflictos semióticos

Estepa y Batanero (1995) identifican una serie de estrategias utilizadas para abordar tareas de asociación estadística de manera inicial; las siguientes son adecuaciones de esas estrategias para una tabla de contingencia 2x2:

- Estratégias correctas (EC):
 - EC1. Comparación de las frecuencias relativas condicionales $h(B|A)$, $h(B|\sim A)$, $h(\sim B|A)$ y $h(\sim B|\sim A)$, indicando de que la dependencia de B respecto a A implicaría la variación de las frecuencias relativas condicionales.
 - EC2. Comparación de las frecuencias relativas $h(B|A)$ y $h(B|\sim A)$ o $h(\sim B|A)$ y $h(\sim B|\sim A)$, haciendo referencia que la independencia entre A y B implica la invarianza de la distribución de B al condicionarla con un valor de A en cada posible valor de B.
 - EC3. Comparación de posibilidades en vez de las probabilidades, es decir, la comparación de la frecuencia de los casos a favor y en contra de B (o sus razones) en cada valor de A argumentando, por ejemplo, indicando que hay una correspondencia unívoca entre la probabilidad de un suceso y la razón entre sus posibilidades a favor y en contra.
- Estratégias parcialmente correctas (EP):
 - EP1. Comparación de las dos frecuencias absolutas con la frecuencia absoluta marginal correspondiente (frecuencia de cada dato al estudiar por separado las variables) mediante comparaciones aditivas o cualitativas, similar a EC1 pero sin el uso explícito de las frecuencias relativas.
 - EP2. Comparación de una única frecuencia absoluta con la correspondiente frecuencia marginal, por ejemplo, indicando que es suficiente la variación de la frecuencia relativa de un solo valor de A cuando B varía para probar la dependencia.
 - EP3. Comparación de la suma de frecuencias en las diagonales, considerada como válida solo si la tabla tiene iguales frecuencias marginales para la variable independiente.
- Estratégias incorrectas (EI):

- EI1. Usar solo una celda de la tabla, generalmente la de frecuencia máxima, y sin usar parte importante de la información, sino solo la sobresaliente (relacionada a la concepción localista de la asociación), en particular, los estudiantes pueden argumentar que no existe dependencia entre las variables porque dos celdas contiguas no tienen frecuencia nula (conocida como concepción determinista de la asociación)
- EI2. Usar solo una distribución condicional $h(B|A)$, es decir, no se concibe el problema como la comparación de probabilidades (se asocia con la concepción localista)
- EI3. Comparar algunas frecuencias dobles con el número total de observaciones, particularmente con la de mayor frecuencia.
- EI4. Comparar frecuencias marginales
- EI5. Otros procedimientos, que aparecen esporádicamente y no se relacionan con las frecuencias de la tabla.

Además, Estepa y Batanero (1995) indican que es posible identificar las siguientes concepciones previas sobre la asociación estadística, algunas de ellas mencionadas anteriormente:

- Concepción causal. Los alumnos solo consideran dependencia entre variables si se le puede atribuir la existencia de una relación del tipo causa-efecto, generalmente se argumenta la inexistencia de una relación con base en que ‘una de ellas no ejerce influencia sobre la otra’. Esta también puede ser extrema, lo que implica que se desechen todas excepto una de las posibles causas que produzcan un mismo efecto.
- Concepción determinista. Significa no admitir excepciones en una relación existente entre las variables. Esta se manifiesta, en una tabla de contingencia, al argumentar que es necesario que una de las frecuencias sea cero para que exista dependencia.
- Concepción unidireccional. Se considera una dependencia solamente cuando es positiva, tomando la asociación inversa como independencia; esto se puede expresar

en el análisis de una tabla de contingencia al indicar que A es independiente de B porque $h(A|B) < h(A|\sim B)$.

- Concepción localista. Corresponde a utilizar solamente parte de los datos de la tabla para emitir un juicio, por ejemplo, leer solo una frecuencia relativa.

Cañadas et al. (2013b) profundizan respecto a las estrategias incorrectas, categorizándolas con base en los siguientes conflictos semióticos:

- Propiedades incorrectas atribuidas a la asociación (PA):
 - PA1. Confundir causalidad con asociación (hace referencia a la concepción causal)
 - PA2. Suponer que la asociación se deduce solo de una parte de los datos (implica concepción local)
 - PA3. Suponer que la asociación se deduce de comparaciones aditivas
 - PA4. Esperar un valor elevado en la primera celda de la tabla.
 - PA5. Concepción determinista.
 - PA6. Concepción unidireccional.
- Propiedades incorrectas atribuidas a la independencia (PI):
 - PI1. Suponer equiprobabilidad en las celdas en caso de independencia, esto que implica confundir independencia con equiprobabilidad.
- Confusión de conceptos (CC): (
 - CC1. Intentar calcular la media de las frecuencias.
 - CC2. Obtener una probabilidad mayor que 1.
 - CC3. Confundir frecuencia doble con marginal (intentando deducir la asociación a partir de las frecuencias marginales).
 - CC4. Confundir razón con frecuencia absoluta (probabilidad con caso posible).
 - CC5. Confundir frecuencia con valor de la variable.
 - CC6. Confundir variable ordinal con cualitativa.

- CC7. Confundir frecuencias dobles con condicionales
- CC8. Confundir los términos de la probabilidad condicional, por ejemplo $P(A|B)$ con $P(B|A)$.
- CC9. Confundir la tabla de contingencia con la de frecuencias.
- Conflictos procedimentales (CP):
 - CP1. Basar el juicio de asociación en teorías previas.
 - CP2. Comparar muestras de distinto tamaño usando frecuencias absolutas.

Resultados de investigaciones que consideraron estas categorías (e.g. Batanero et al., 1998; Cañadas et al., 2013b) para el estudio de los juicios de asociación, muestran buenas bases intuitivas en estudiantes de Educación Superior; sin embargo, también evidencian la presencia de los errores y las concepciones mencionadas anteriormente. Con base en esto, es posible hipotetizar que dichos errores y concepciones mencionadas se presenten en las respuestas de los estudiantes de Educación Media, y que estas muestren una base intuitiva más débil en comparación con las respuestas de estudiantes de Educación Superior.

Metodología

Esta investigación se enmarca en una metodología de corte cualitativo (Pérez-Serrano, 1994), con un nivel exploratorio-descriptivo (Hernández, Fernández y Baptista, 2010), debido a que se centra en explorar las respuestas de los estudiantes, con el propósito de describir las estrategias y conflictos que presentan al abordar tareas de juicios de asociación sin previa instrucción.

Participantes

La muestra estudiada, de tipo no probabilística seleccionada por conveniencia, corresponde a 75 estudiantes pertenecientes a Educación Media (25 de 1° Medio, 22 de 2° Medio, 14 de 3° Medio y 14 de 4° Medio) de un establecimiento particular subvencionado de la ciudad de Osorno de Chile, cuyas oscilaban de 14 a 19 años. Aunque los estudiantes no hicieron un trabajo previo respecto al tema (métodos directos para su cálculo, por

ejemplo, el coeficiente Phi de Pearson) ya habían abordado conceptos base para realizar un acercamiento inicial a las tareas (frecuencia, tablas de contingencia, entre otros).

Instrumento de indagación y su aplicación

El instrumento de indagación corresponde un cuestionario con tres tareas sobre juicios de asociación en tablas de contingencia, sobre situaciones problemas diferentes. Dicho cuestionario de valido por medio de una triangulación de contenidos, una triangulación de expertos, aplicación del cuestionario piloto y una validación interjueces. Este fue aplicado en una sesión de hora y media con la supervisión del profesor encargado. Las situaciones problemas, así como su tarea respectiva, se presentan en las Figuras 1, 2 y 3:

Problema 1. Se quiere estudiar si el fumar produce molestias respiratorias. Para esto, se ha observado a un grupo de 400 personas durante un periodo suficiente de tiempo, obteniendo los siguientes resultados:

	Tiene molestias respiratorias	No tiene molestias respiratorias
Fuma	180	40
No fuma	20	160

¿Hay relación entre el fumar (Fuma o No fuma) y el tener molestias respiratorias (Tiene molestias respiratorias o No tiene molestias respiratorias)? ¿O son las variables independientes? Explica tu respuesta.

Figura 1 - Situación problema 1: Molestias respiratorias y fumadores

Fuente: elaborado por los autores (Adaptado de Estrada y Díaz, 2007)

Problema 2. Se quiere estudiar si hay relación entre la edad de una persona y el tener un ataque al corazón. Para ello, en un centro médico han sido observados 200 personas. Los resultados son los siguientes:

	Menor o igual de 55 años	Mayores de 55 años
Ha tenido un ataque al corazón	20	80
Nunca ha tenido un ataque al corazón	90	10

¿Hay relación entre el tener un ataque al corazón (Ha tenido un ataque al corazón o Nunca ha tenido un ataque al corazón) y la edad de una persona (Menor o igual de 55 años, o Mayores de 55 años)? ¿O son las variables independientes? Explica tu respuesta.

Figura 2 - Situación problema 1: Mayores de 55 años y ataques al corazón

Fuente: elaborado por los autores (Adaptado de Estrada y Díaz, 2006)

Problema 3. Se quiere estudiar si hay relación entre el género de los alumnos y el gusto por el ping-pong. Para esto, en un colegio se pregunta a 180 alumnos si les gusta o no el ping-pong, obteniendo los siguientes resultados:

	Chicos	Chicas
Le gusta el ping-pong	80	40
No le gusta el ping-pong	40	20

¿Hay relación entre el género de los alumnos (Chicos o Chicas) y el gusto por el ping-pong (Le gusta el ping-pong o No le gusta el ping-pong)? ¿O son las variables independientes? Explica tu respuesta.

Figura 3 - Situación problema 3: Género y gusto por el Ping-Pong

Fuente: elaborado por los autores (Adaptado de Contreras, Estrada, Batanero y Díaz, 2010)

Cabe señalar que, con base en el coeficiente Phi de Pearson, las variables de la situación 1 presentan una correlación positiva ($r = 0,7035$), las de la situación 2 presentan una correlación negativa ($r = -0,7035$), y las de la situación 3 no presentan correlación ($r = 0$).

Procedimiento de análisis de datos recolectados

El proceso de análisis consistió en dos etapas: (1) identificar las estrategias y concepciones presentes en las respuestas de los estudiantes, tomando como referencia las señaladas en los fundamentos del estudio, y (2) presentar los resultados en tablas de frecuencias, organizados por situación trabajada y año escolar, con el objetivo de identificar aspectos relevantes respecto a cómo los estudiantes abordan de manera inicial este tipo de tarea. Para que una respuesta se clasifique en una de las categorías indicadas por Estepa y Batanero (1995), y en consideración a que se aplicó el instrumento sin previa instrucción, se toma como condición suficiente el que los estudiantes hagan referencia a la estrategia respectiva mencionando probabilidades, posibilidades, comparación de frecuencias, entre otros.

Resultados



En las respuestas de los estudiantes fue posible de identificar los tres tipos de estrategias (correcta, parcialmente correcta e incorrecta), además de casos en los que no se responde la pregunta (S/R) o no proporcionan la justificación (S/J).

De los estudiantes que hicieron referencias a estrategias correctas (EC), fue posible identificar respuestas pertenecientes a las tres subcategorías (EC1, EC2 y EC3). Por ejemplo, la Figura 4 corresponde a la respuesta de un estudiante que indica que quienes fuman tienen mayor probabilidad de tener molestias respiratorias (hace referencia a $h(A|B)$) y que solo el 10% de los que no fuman padecen molestias respiratorias $h(A|\sim B)$, es decir, haciendo referencia a las probabilidades condicionales de un valor de A específico para cada valor de B (EC2).

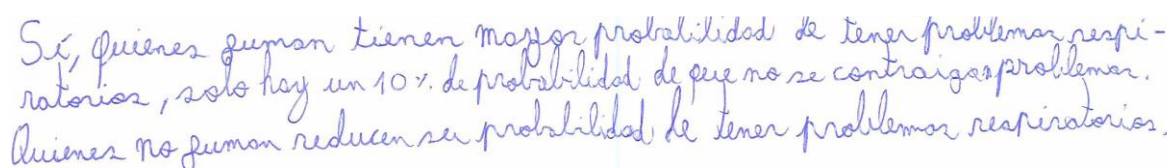


Figura 4 - Respuesta tipo EC2 a la tarea de la situación problema 1 (Estudiante de 4° Medio)

Fuente: información recolectada por los autores para la investigación

Asimismo, se identificaron respuestas que mostraban estrategias parcialmente correctas, clasificándolas en las 3 subcategorías (EP1, EP2 y EP3). En la Figura 5 se presenta la respuesta de un estudiante que argumentó con base en lo señalado en la categoría EP2, en concreto, realizando una comparación entre las frecuencias de $(\sim A \cap B)$ y $(A \cap B)$.

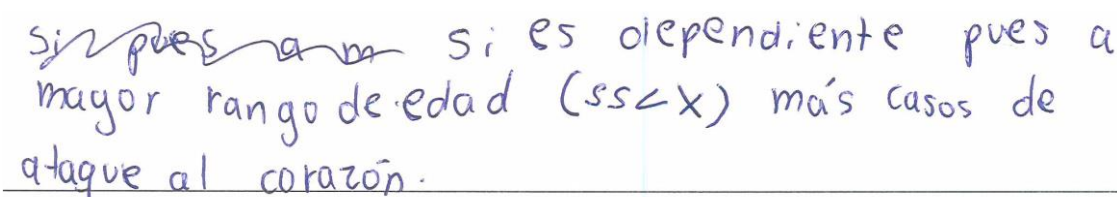


Figura 5 - Respuesta tipo EP2 a la tarea de la situación problema 2 (Estudiante de 4° Medio)

Fuente: información recolectada por los autores para la investigación

Con respecto a las respuestas que presentan estrategias incorrectas, fue interesante observar que la mayor proporción de los estudiantes correspondió a la argumentación con base en la concepción causal o la concepción determinista, es decir, atribuyendo la existencia de una dependencia a identificar una relación del tipo causa-efecto, por ejemplo, estableciendo una relación causa-efecto entre la edad de la persona y el hecho de haber sufrido un ataque al corazón (Figura 6); o bien, al no admitir excepciones en una relación existente entre las variables, por ejemplo, argumentado independencia como implicancia de que hay casos en los que no se cumple la relación (Figura 7).

son dependientes por que las personas con más edad son más propensas a esto.

Figura 6 - Respuesta tipo PA1 a la tarea de la situación problema 1 (Estudiante de 1° Medio)

Fuente: Fuente: información recolectada por los autores para la investigación

no hay relación yo que a cada uno le puede gustar lo que quiere.

Figura 7 - Respuesta tipo PA5 a la tarea de la situación problema 3 (Estudiante de 1° Medio)

Fuente: información recolecta por los autores para la investigación

La Tabla 2 sintetiza los resultados obtenidos de las respuestas con respecto a la identificación de dependencia o independencia de las variables, así como la estrategia (correcta, parcialmente correcta o incorrecta) utilizada para abordar el problema.

Tabla 2. Frecuencia de respuesta dada y estrategia utilizada

CURSO	SITUACIÓN PROBLEMA	¿EXISTE DEPENDENCIA?	EC	EP	EI	S/J
-------	--------------------	----------------------	----	----	----	-----

		SI	NO	S/R	1	2	3	1	2	1	2	4	5	
1° M	P1	12	2	11	1	0	1	1	0	0	0	0	8	3
	P2	11	3	11	0	0	1	1	1	0	0	0	9	2
	P3	3	11	11	0	0	0	2	0	0	0	1	7	4
2° M	P1	9	2	11	0	4	0	2	0	0	0	0	5	0
	P2	8	4	10	0	3	0	2	0	0	0	0	2	5
	P3	2	7	13	0	0	0	1	0	0	0	0	5	3
3° M	P1	8	6	0	0	1	0	0	2	0	0	0	8	3
	P2	7	7	0	0	1	0	0	1	1	0	0	5	6
	P3	5	6	3	0	1	0	0	1	0	0	0	0	9
4° M	P1	11	1	2	0	1	2	1	4	1	0	0	0	3
	P2	13	1	0	0	0	0	0	10	0	2	0	0	2
	P3	5	7	2	0	0	1	1	1	0	0	0	2	7

Fuente: elaborada por los autores

Con respecto a la identificación de dependencia o independencia, es posible señalar que, independientemente del curso, la mayoría de los estudiantes que responden la pregunta identifican correctamente la dependencia o independencia. Esto es más notorio en 1° Medio (86% de quienes respondieron P1 identificaron dependencia, 79% dependencia en P2 y 79% independencia en P3) y 4° Medio (92% de los estudiantes que respondieron P1 y un 93% que respondieron P2 identificaron la dependencia).

Por otro lado, en 1° Medio y 2° Medio, una gran cantidad de estudiantes opta por no responder el problema o no entrega una respuesta clara; específicamente, un 44% para P1, P2 y P3 en 1° Medio, y un 50%, 45% y 59% para P1, P2 y P3 en 2° Medio, respectivamente. Esto contrasta con la cantidad de estudiantes de 3° Medio y 4° Medio que no responden los problemas, en particular, el 21% de los estudiantes de 3° Medio no respondió P3, y de 4° Medio, el 14% no respondió P1 y P3.

Respecto a las estrategias utilizadas, es posible identificar que los estudiantes no logran relacionar correctamente los juicios de asociación con las probabilidades condicionales o frecuencias relativas, lo que se observa en la máxima cantidad de estrategias correctas (EC) y parcialmente correctas (EP) por pregunta, de cada curso: 1° Medio: 14% de EC para P1 y 14% de EP para P2 y P3; 2° Medio: 36% de EC para P1 y 18% de EP para P1; 3° Medio: 7% de EC para P1, 9% para P3 y 14% de EP para P1; 4° Medio: 25% EC para P1 y 71% EP para P2. Además, es posible identificar un uso de estrategias incorrectas por aquellos estudiantes que responden los problemas, especialmente en 1° Medio (57% para P1, 64% para P2 y 57% para P3), así como la falta de justificación, que se presenta con mayor frecuencia para P3.

La identificación de conflictos semióticos en las respuestas de los estudiantes reflejó un panorama muy favorable, puesto que la totalidad de estos corresponden a la categoría de propiedades incorrectas atribuidas a la asociación (PA), es decir, no se presentaron conflictos respecto a las propiedades de la independencia, la confusión entre conceptos esenciales para realizar juicios de asociación (e.g. frecuencia doble con frecuencia marginal) o la confusión respecto al procedimiento realizado (ver Tabla 3). Esto sugiere que los estudiantes de Educación Media que participaron en este estudio no presentan mayores dificultades en lo que involucra a saberes previos necesarios para el aprendizaje de los juicios de asociación en tablas de contingencia 2x2, sino en el concepto.

Tabla 3. Frecuencia de conflictos identificados en las respuestas

CURSO	SITUACIÓN PROBLEMA	PA			
		1	2	5	6
1° M	1	4	0	0	0
	2	7	0	2	0
	3	0	0	5	0
2° M	1	2	0	0	1

	2	1	0	0	0
	3	1	0	2	0
3° M	1	2	1	2	0
	2	1	0	3	0
	3	0	0	0	0
4° M	1	1	0	0	0
	2	2	0	0	0
	3	0	2	0	0

Fuente: elaborada por los autores

De los conflictos identificados, la mayoría (54% del total) corresponde a concepciones del tipo causal (indicando una dependencia o independencia entre las variables con base en una relación causa-evento), seguida por concepciones deterministas (36% del total de conflictos). Por otro lado, es posible identificar que en 4° Medio se presenta la menor cantidad de conflictos (13% del total), y no se presentan los conflictos de asumir que la dependencia o independencia se deduce de comparaciones aditivas (PA3), o de esperar un valor elevado en la primera celda de la tabla (PA4).

Conclusiones

Los resultados sugieren que, en general, los estudiantes poseen una buena base intuitiva para realizar juicios de asociación, lo que coincide con Cañadas et al. (2013b). Respecto a las estrategias utilizadas, se puede indicar que los estudiantes no justifican sus juicios de asociación con las probabilidades condicionales o frecuencias relativas condicionales, lo que se observa en los bajos porcentajes de estrategias correctas (EC) y parcialmente correctas (EP), y en el alto porcentaje de estrategias incorrectas, especialmente en 1° Medio. Esto sugiere que los estudiantes de 1° Medio tienen dificultad para justificar sus respuestas con base en principios probabilísticos o matemáticos y que, principalmente en los primeros años, tienen una mayor dificultad para justificar la independencia entre variables, a pesar de que puedan identificarla intuitivamente.

Por otro lado, el problema 1 fue el que presento el mayor porcentaje de estrategias correctas (EC). Esto sugiere que, en ese problema, a los estudiantes les fue más fácil asociar las frecuencias relativas condicionales y la probabilidad condicional con el juicio de asociación, posiblemente debido a que se presentan frecuencias con mayor diferencia, lo que les facilita realizar comparaciones.

Con respecto a los conflictos semióticos identificados, los resultados evidencian que los estudiantes de Educación Media presentan dificultades para diferenciar una relación de causa-efecto con una relación de dependencia. Además, la poca presencia de este tipo de conflictos en cursos superiores (2°, 3° y 4° Medio) sugieren que, posiblemente, los estudiantes adquieren nociones sobre esta diferencia al cursar el contenido de tablas de contingencia en la Unidad 4 ‘Probabilidad y Estadística’ en 1° Medio (MINEDUC, 2016).

Con este estudio identificamos diversos desafíos para el aprendizaje del concepto de asociación entre dos variables cualitativas aplicado en tablas de contingencia, principalmente, superar la concepción causal, el desarrollo del razonamiento intuitivo en estudiantes de 1° Medio y 2° Medio, y la argumentación. Desafíos que podrían ser abordados a través de una trayectoria hipotética de aprendizaje, o mediante el diseño de actividades de aprendizaje.

Referencias

- Batanero, C., Cañadas, G. R., Contreras, J. M., y Gea, M. M. (2015). La comprensión de las tablas de contingencia: una síntesis de la investigación didáctica. *Boletín de Estadística e Investigación Operativa*, 31(3), 299-315. Recuperado de: <http://www.ugr.es/~batanero/documentos/BEIOTabla.pdf>
- Batanero, C., Godino, J., y Estepa, A. (1998). La construcción del significado de la asociación mediante actividades de análisis de datos: Reflexiones sobre el papel del ordenador en la enseñanza de la estadística. En J. R. Pascual (Ed.), *II Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 169-185). Pamplona, España: SEIEM. Recuperado de: http://funes.uniandes.edu.co/1494/1/Batanero1999La_SEIEM_169.pdf

- Cañadas, G. R., Batanero, C., Estepa, A., y Arteaga, P. (2013b). Juicios de asociación en tablas de contingencia con datos ordinales. En A. Berciano, G. Gutiérrez, A. Estepa y N. Climent (Eds.), *Investigación en Educación Matemática XVII* (pp. 209-218). Bilbao: SEIEM. Recuperado de: <http://funes.uniandes.edu.co/6218/1/Ca%C3%B1adas2013JuiciosSEIEM.pdf>
- Cañadas, G. R., Contreras, J. M., Arteaga, P., y Gea, M. M. (2013a). Problemática y recursos en la interpretación de las tablas de contingencia. *Unión. Revista Iberoamericana de Educación Matemática*, 34, 85-96. Recuperado de: <http://www.ugr.es/~jmcontreras/pages/Investigacion/articulos/2013Union2.pdf>
- Cañadas, G., Arteaga, P., Guirado, R. y Molina, E. (2017). Conflictos semióticos en el cálculo de probabilidades condicionales y juicios de asociación en una tabla de contingencia. En J. M. Contreras, P. Arteaga, G. R. Cañadas, M. M. Gea, B. Giacomone y M. M. López-Martín (Eds.), *Actas del Segundo Congreso Internacional Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos*. Recuperado de: <https://digibug.ugr.es/bitstream/handle/10481/45417/canadas.pdf?sequence=1&isAlloved=y>
- Cañadas, G., Contreras, J.M., Gea, M.M., y Arteaga, P. (2012). Comprensión de frecuencias y representaciones asociadas a las tablas de contingencia. En F. España y B. Sepúlveda (Eds.), *XIV Congreso de Educación y Aprendizaje Matemático* (pp. 129-138). Málaga, España: S.A.E.M. THALES. Recuperado de: <http://funes.uniandes.edu.co/21666/1/Ca%C3%B1adas2012Comprension.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.M. Moreno, A. Estrada, J. Carrillo y T. A. Sierra, (Eds.), *Investigación en Educación Matemática XIV* (pp. 271-280). Lleida: SEIEM. Recuperado de: <https://core.ac.uk/download/pdf/12342072.pdf>
- Esquivel, E. C. (2016). La enseñanza de la Estadística y la Probabilidad, más allá de procedimientos y técnicas. *Cuadernos de Investigación y Formación en Educación Matemática*, 11(15), 21-31. Recuperado de: <https://revistas.ucr.ac.cr/index.php/cifem/article/view/23880/24056>
- Estepa, A. (2007). Caracterización del significado de la correlación y regresión en estudiantes de Educación Secundaria. *Zetetiké*, 15(2), 119-152. Recuperado de: <https://periodicos.sbu.unicamp.br/ojs/index.php/zetetike/article/view/8647028/13929>

- Estepa, A. y Batanero, C. (1995). Concepciones iniciales sobre la asociación estadística. *Enseñanza de las ciencias*, 13(2), 155-170. Recuperado de: <https://www.raco.cat/index.php/Ensenanza/article/download/21404/93364>
- Estrada, A. y Díaz, C. (2006). Computing probabilities from two-way tables. An exploratory study with future teachers. En Rossman, A., Chance, B. (Eds.), *Seventh International Conference on Teaching of Statistics*. Salvador (Bahia): International Association for Statistical Education. Recuperado de: <https://www.stat.auckland.ac.nz/~iase/publications/17/C413.pdf>
- Estrada, M. y Díaz, C. (2007). Errores en el cálculo de probabilidades en tablas de doble entrada en profesores en formación. *Uno. Revista de Didáctica de las Matemáticas*, 44, 48-57. Recuperado de: <https://core.ac.uk/download/pdf/70291152.pdf>
- Gea, M. M., Batanero, C., López-Martín, M., y Arteaga, P (2016). Investigaciones sobre la percepción y el aprendizaje de la correlación y regresión. *Boletín de Estadística e Investigación Operativa*, 32(3), 184-206.
- Hernández, R., Fernández, C., y Baptista, P. (2010). Metodología de la investigación. México: McGraw-Hill.
- Lavalle, A., Micheli, E., y Rubio, N. (2006). Análisis didáctico de regresión y correlación para la enseñanza media. *Revista Latinoamericana de Investigación en Matemática Educativa*, 9(3), 383-406. Recuperado de: <http://www.scielo.org.mx/pdf/relime/v9n3/v9n3a4.pdf>
- MINEDUC (2016). *Programa de Estudio Primero Medio*. Santiago: Ministerio de Educación. Recuperado de: https://www.curriculumnacional.cl/614/articulos-34359_programa.pdf
- Pérez-Serrano, G. (1994). *Investigación cualitativa. Retos e interrogantes. I. Métodos*. Madrid: La Muralla.
- Vásquez, C., Ruz, F., y Martínez, M. (2020). Recursos virtuales para la enseñanza de la estadística y la probabilidad: un aporte para la priorización curricular chilena frente a la pandemia de la COVID-19. *TANGRAM-Revista de Educação Matemática*. 3(2), 159-183. Recuperado de: <https://ojs.ufgd.edu.br/index.php/tangram/article/view/12299/5855>

Contribución de los autores

1º autor: define las variables a analizar, genera la metodología, analiza los datos recolectados, realiza el escrito del artículo, revisa la versión final del artículo.

2º autor: gestiona la idea, entrega antecedentes, analiza los datos recolectados, colabora con la redacción, revisa la versión final del artículo.

3º autor: implementa el instrumento, recoge los datos, colabora en la revisión de los datos, aporta modificaciones al escrito.

4º autor: entrega antecedentes, revisa el artículo generando propuestas de modificación, revisa la versión final del artículo.