

Tablas de doble entrada: lectura y cálculo de probabilidades por estudiantes de educación media

*Daniela Calderón Torres*¹

*Jaime I. García-García*²

*Nicolás Fernández Coronado*³

*Elizabeth H. Arredondo*⁴

RESUMEN

La tabla de doble entrada es una representación útil para el análisis de dos variables estadísticas, cuya enseñanza es esencial en la formación del estudiante. Esta investigación cualitativa, descriptiva-exploratoria, tuvo por objetivo analizar la lectura y el cálculo de probabilidades en tablas de doble entrada realizado por estudiantes de educación media. Bajo la técnica de análisis de contenido, se analizan las respuestas de 75 estudiantes a dos tareas de lectura y tres de cálculo de probabilidades. Los resultados muestran que la mayoría del estudiantado posee dominio de los primeros niveles de lectura de Curcio y hace uso de la regla de Laplace, o la regla de tres, como estrategia de solución; sin embargo, algunos presentan errores, por ejemplo, confunden una probabilidad condicional con su inversa. Lo anterior evidencia la necesidad de diseñar propuestas de enseñanza donde se aborden los diversos objetos y procesos matemáticos en torno a la lectura y cálculo de probabilidades en tablas de doble entrada.

PALABRAS CLAVE: Estrategia. Error. Nivel de lectura. Cálculo de probabilidades. Tablas de contingencia.

¹ Profesora en Educación Media, mención Matemática y Computación. Universidad de Los Lagos, Osorno, Chile. Orcid: <https://orcid.org/0000-0001-7036-3454>. E-mail: danieladelapilar.calderon2@alumnos.ulagos.cl.

² Doctor en Ciencias, especialidad en Matemática Educativa. Universidad Metropolitana de Ciencias de la Educación, Santiago, Chile. Orcid: <https://orcid.org/0000-0002-8799-5981>. E-mail: jaime.garcia@umce.cl.

³ Estudiante del Magíster en Educación Matemática. Universidad de Los Lagos, Osorno, Chile. Orcid: <https://orcid.org/0000-0002-9613-3144>. E-mail: nicolasalonso.fernandez@alumnos.ulagos.cl.

⁴ Doctora en Ciencias, especialidad en Matemática Educativa. Universidad de Los Lagos, Osorno, Chile. Orcid: <https://orcid.org/0000-0002-5285-1603>. E-mail: elizabeth.hernandez@ulagos.cl.

Tabelas de dupla entrada: leitura e cálculo de probabilidades por alunos do ensino médio

RESUMO

A tabela de dupla entrada é uma representação útil para a análise de duas variáveis estatísticas, cujo ensino é essencial na formação do aluno. Esta investigação qualitativa, de carácter descritivo-exploratório, teve como objetivo analisar a leitura e o cálculo de probabilidades em tabelas de dupla entrada efetuados por alunos do ensino médio. Utilizando a técnica de análise de conteúdo, foram analisadas as respostas de 75 alunos a duas tarefas de leitura e a três tarefas de cálculo de probabilidades. Os resultados mostram que a maioria dos alunos domina os primeiros níveis de leitura de Curcio e utiliza a regra de Laplace, ou a regra de três, como estratégia de solução; no entanto, alguns cometem erros, por exemplo, confundem uma probabilidade condicional com a sua inversa. Isto mostra a necessidade de conceber propostas de ensino que abordem os diferentes objetos e processos matemáticos em torno da leitura e cálculo de probabilidades em tabelas de dupla entrada.

PALAVRAS-CHAVE: Estratégia. Erro. Nível de leitura. Cálculo de probabilidades. Tabelas de contingência.

Double-entry tables: reading and calculating probabilities by high school students

ABSTRACT

The double-entry table is a useful representation for the analysis of two statistical variables, whose teaching is essential in the student's education. This qualitative research, descriptive-exploratory, aimed to analyze the reading and calculation of probabilities in double-entry tables carried out by high school students. Using the content analysis technique, the responses of 75 students to two reading and three probability calculation tasks are analyzed. The results show that the majority of the students have mastery of the first levels of Curcio's reading and make use of Laplace's rule, or the rule of three, as a solution strategy; however, some have errors, for example, they confuse a conditional probability with its inverse. The above evidence the need to

design teaching proposals that address the various mathematical objects and processes related to the reading and calculation of probabilities in double-entry tables.

KEYWORDS: Strategy. Error. Reading levels. Probability calculation. Contingency tables.

* * *

Introducción y antecedentes

En los últimos años, la enseñanza de la probabilidad y la estadística ha tomado relevancia debido a que coadyuvan con conocimientos y habilidades a los estudiantes para interpretar mensajes probabilísticos, comprender información estadística representada en tablas o gráficos, comprobar conjeturas y tomar decisiones basadas en el análisis de datos; es decir, les permiten hacer frente y encontrar soluciones a situaciones cotidianas donde la aleatoriedad y la variabilidad están presentes.

En didáctica de la probabilidad y la estadística, desde hace dos décadas se han estado desarrollando estudios enfocados en los términos conocidos como alfabetización estadística y alfabetización probabilística (GAL, 2002; 2005). Estos términos consisten en: 1) la capacidad de interpretar y evaluar críticamente información estadística o probabilística de diversos contextos y, con ello, tomar decisiones fundamentadas; y 2) la capacidad para formular, discutir o comunicar opiniones con respecto a dicha información. Diversos autores (GAL, 2019; RODRÍGUEZ-ALVEAL et al., 2018) señalan ambas alfabetizaciones como necesidades relevantes en la educación de los estudiantes y en la formación de profesores, con el propósito de formar ciudadanos alfabetizados estadística y probabilísticamente.

En Gal (2005) se señala como elemento esencial de la alfabetización probabilística el cálculo de probabilidades, refiriéndose a las formas para calcular o estimar la probabilidad de eventos; siendo los enfoques clásico, frecuencial y subjetivo de la probabilidad, aquellos que se promueven en el

currículo escolar de matemáticas de diversos países (tales como Chile y México) para el cálculo o estimación de probabilidades (GAL, 2005; SÁNCHEZ, 2009).

Con respecto a la tabla estadística, ésta es útil para la organización, descripción y análisis de datos. Schield (2006) señala que una persona alfabetizada estadísticamente debiera leer críticamente las tablas estadísticas que comúnmente se encuentra en su entorno laboral, en la prensa, etcétera; pero no solo leerlas de manera literal, sino identificar la variabilidad y tendencia en los datos, y los posibles errores que puedan distorsionar la información que se pretende comunicar.

La tabla de doble entrada es una representación utilizada para registrar y presentar la distribución de frecuencias de una variable estadística bidimensional. Su utilidad para el estudio de datos cualitativos, su presencia en los medios de difusión de información y el mundo laboral, así como su uso en tareas relacionadas con el cálculo de probabilidades, vuelven su lectura e interpretación como elementos básicos de la alfabetización estadística y probabilística (ESTRADA; DÍAZ, 2007); bajo esta perspectiva, se considera que su enseñanza es esencial en la formación del estudiante. Sin embargo, diversos autores identifican niveles básicos respecto a la comprensión de tablas estadísticas y errores en el cálculo de probabilidades en tablas de doble entrada. A continuación, se mencionan algunos estudios relacionados con esta temática.

En relación con la lectura de tablas estadísticas, Rodríguez y Sandoval (2012) analizan la lectura y construcción de gráficos y tablas por 47 profesores chilenos de educación básica en servicio y 44 en formación, identificando que ambos grupos presentan habilidades básicas de lectura, es decir, se ubican en el nivel de lectura 1 de Curcio, leer los datos, y que solo algunos participantes alcanzan el nivel 2, leer dentro de los datos. Díaz-Levicoy et al. (2016) analizan el nivel de lectura de tablas estadísticas que alcanzan 121 profesoras chilenas en formación de educación infantil, encontrando que la mayoría de las participantes dominan los niveles 1 y 2, leer los datos y leer dentro de los datos; sin embargo, solo la tercera parte de las participantes alcanza el nivel 3, leer

más allá de los datos, al predecir algún dato o tendencia a partir de la información entregada en la tabla. Por su parte, García-García et al. (2019) realizan un análisis comparativo sobre el nivel de comprensión de una tabla estadística entre estudiantes universitarios mexicanos y chilenos (36 y 35, respectivamente), considerando una jerarquía propuesta a partir de la articulación de los niveles de Curcio y los niveles de Aoyama; sus resultados muestran que la mayoría de los estudiantes de ambos países alcanzan el nivel 2, comparativo; sin embargo, los estudiantes mexicanos presentan mayores errores en la comparación de los datos. Gea et al. (2020) analizan las respuestas de 69 futuros profesores de educación primaria a una tarea que implica la construcción y lectura de una tabla de doble entrada; con respecto a la lectura, los participantes manifiestan dificultades al identificar frecuencias conjuntas y condicionales que implican cálculos de probabilidades.

Respecto al cálculo de probabilidades en tablas de doble entrada, Díaz y De la Fuente (2005) analizan las dificultades que presentan 154 psicólogos en formación, identificando que, aunque la mayoría de ellos calcula correctamente la probabilidad simple, pocos calculan la probabilidad condicional y compuesta, identificando que los participantes confunden un suceso con su complemento y la probabilidad con casos favorables. Estrada y Díaz (2006; 2007) analizan los conflictos semióticos que presentan futuros profesores al calcular probabilidades en tablas de doble entrada, con resultados similares a Díaz y De la Fuente (2005) en cuanto al manejo de las probabilidades, e identificando conflictos semióticos como confundir una probabilidad condicional con su inversa, confundir una probabilidad condicional con una conjunta y confundir un suceso con su complemento. Por su parte, Contreras et al. (2010) analizan las respuestas de 69 profesores en formación de educación primaria con respecto al cálculo de probabilidades (simple, compuesta y condicional) en tablas de doble entrada. En comparación con los resultados de Estrada y Díaz (2006), se presenta un bajo porcentaje de respuestas correctas y una mayor proporción de participantes que no abordan las tareas solicitadas, mientras

que, entre los conflictos semióticos reportados, se destaca la confusión de la condicional con su inversa.

Los resultados de estos estudios, realizados con estudiantes universitarios y profesores, evidencian bajos niveles de lectura y errores presentes en el cálculo de probabilidades, que corresponden a desafíos en la formación de estudiantes alfabetizados estadística y probabilísticamente. Bajo esa perspectiva, y con el propósito de explorar esta línea investigativa en otro nivel educativo y presentar un panorama que facilite abordar los errores desde niveles menores de educación, nos planteamos como objetivo analizar la lectura y el cálculo de probabilidades en tablas de doble entrada realizado por estudiantes chilenos de educación media.

Análisis semiótico de la tabla de doble entrada

La tabla de doble entrada es un cuadro, estructurado por columnas y filas, cuya finalidad es resumir información acerca de dos variables estadísticas y registrar la distribución conjunta de dichas variables (GEA et al., 2020). Contiene tantas filas y columnas como categorías que presenten las dos variables estadísticas que la constituyen. Debido a que en esta representación subyacen varios conceptos y sus interrelaciones, la tabla de doble entrada es considerada como un objeto semiótico complejo (GEA et al., 2020), cuya forma más simple es cuando las variables poseen solo dos categorías (ESTRADA; DÍAZ, 2007) (ver Tabla 1).

TABLA 1: Formato de la forma más simple de una tabla de doble entrada

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

Fuente: Adaptada de Estrada y Díaz (2007).

De esta tabla, es posible derivar los siguientes tipos de frecuencias:

- Frecuencias absolutas marginales: Cantidades en la columna derecha y en la fila inferior, por tanto, pueden ser por filas ($a+b$ y $c+d$) y por columnas ($a+c$ y $b+d$).
- Frecuencias absolutas dobles: Cantidades de las cuatro celdas centrales (a , b , c y d) e indican la frecuencia en la que se presentan intersecciones específicas de valores de las variables.
- Frecuencias absolutas condicionales: Corresponden a la frecuencia para el valor de una variable, dejando fijo un valor de la otra. Matemáticamente, estas frecuencias absolutas condicionales son iguales a las dobles; sin embargo, no se perciben psicológicamente de la misma manera, debido a que su lectura atiende a la condición indicada, es decir, con relación a distintos valores (la cantidad total de la que forma parte no es la misma) (GEA et al., 2020).

De estas frecuencias absolutas se pueden determinar las frecuencias relativas marginales, dobles y condicionales (no coincidiendo las frecuencias relativas dobles con las relativas condicionales); y asumiendo equiprobabilidad de todos los casos en la muestra, se pueden calcular las probabilidades asociadas (ESTRADA; DÍAZ, 2007; GEA et al., 2020).

Taxonomía de Curcio

Para analizar la lectura de las tablas de doble entrada realizada por los estudiantes de educación media, se consideró la taxonomía de Curcio (CURCIO, 1989; FRIEL; CURCIO; BRIGHT, 2001). Esta taxonomía se estableció para la lectura de gráficos estadísticos; sin embargo, ha sido adaptada y utilizada para el análisis de la lectura de tablas estadísticas (DÍAZ-LEVICOY et al., 2016; GARCÍA-GARCÍA et al., 2019):

- Nivel 1. Leer los datos: corresponde a la lectura literal de datos de la tabla de doble entrada, sin realizar su interpretación ni cálculos adicionales.
- Nivel 2. Leer dentro de los datos: corresponde a la interpretación e integración de los datos de la tabla de doble entrada, lo que implica comparar datos o aplicar algún cálculo matemático sencillo.
- Nivel 3. Leer más allá de los datos: corresponde a la realización de alguna inferencia o predicción a partir de los datos, sobre información no explícita en la tabla de doble entrada.
- Nivel 4. Leer detrás de los datos: corresponde a la valoración crítica de la información representada en la tabla de doble entrada (validez y fiabilidad), la recolección y organización de los datos, las interpretaciones o conclusiones realizadas por otra persona, entre otras consideraciones.

Esta investigación se enfoca en el análisis del dominio de los niveles 1 y 2 de lectura de los estudiantes de educación media; estos niveles están relacionados, respectivamente, con la lectura de la frecuencia absoluta doble y marginal (cuando no se muestra en la tabla de doble entrada). Esto porque consideramos ambos niveles como esenciales para el cálculo de probabilidades en este tipo de representación estadística.

Error y estrategia

El estudio de los errores es un foco de interés en la investigación en educación matemática, debido a que ayuda a explicar parte de la problemática del aprendizaje de las matemáticas (RICO, 1996; SOCAS, 1997). Según Socas (1997), un error se genera por dificultades y obstáculos en el proceso pedagógico y un esquema cognitivo inadecuado en el estudiante, y se entiende como la manifestación externa (escrita, verbal, entre otras) de estos, por ejemplo, en el uso de símbolos o algoritmos matemáticos.

Por otro lado, de acuerdo con Poggioli (1999), una estrategia para resolver problemas se entiende como la operación mental utilizada por el estudiante para

pensar sobre la representación de las metas y los datos, con el propósito de transformarlos y obtener una solución; esta estrategia incluye los métodos heurísticos, los algoritmos y los procesos de pensamiento divergente.

Bajo esta perspectiva, en esta investigación nos enfocamos en el análisis de los errores y las estrategias que manifiestan los estudiantes de educación media al resolver tareas que involucran el cálculo de probabilidades simples, conjuntas y condicionales en tablas de doble entrada; consideramos que lo anterior podría contribuir para el diseño de experiencias de aprendizaje o secuencias didácticas que permitan superar tales errores.

Metodología

Esta investigación sigue una metodología de tipo cualitativa, de nivel descriptivo-exploratorio (HERNÁNDEZ; FERNÁNDEZ; BAPTISTA, 2010). Bajo la técnica de análisis de contenido (KRIPPENDORFF, 1997), se analizan las respuestas proporcionadas por estudiantes a tareas relacionadas con la lectura y cálculo de probabilidades en tablas de doble entrada. En concreto, se identifica el dominio o no dominio de los niveles 1 y 2 de lectura de Curcio, así como las estrategias y los errores en el cálculo de probabilidades simples, conjuntas y condicionales.

La muestra, de tipo no probabilística, estuvo conformada por 75 estudiantes de educación media (25, 22, 14 y 14 estudiantes de 1°, 2°, 3° y 4° medio, respectivamente), seleccionados por conveniencia, cuyas edades oscilaban entre los 14 y 19 años. En Chile, en primer año medio se abordan de manera conjunta las tablas de doble entrada y el cálculo de probabilidades, es decir, se presenta el cálculo de probabilidades en tablas de doble entrada, lo que implica la lectura de este tipo de representaciones estadísticas (MINEDUC, 2016). Cabe señalar que, al momento de realizar la investigación, los estudiantes habían recibido enseñanza sobre este tema (esta información se corroboró con los profesores que imparten el curso de matemáticas en primero medio).

Como instrumento, se diseñó un cuestionario con cinco tareas relacionadas con un problema contextualizado (ver Figura 1) que trata la relación entre dos variables dicotómicas: edad y registro de algún ataque al corazón. Al seleccionar una persona al azar se realiza un fenómeno compuesto por dos experimentos aleatorios: el primero con dos sucesos: $A =$ “menor o igual de 55 años” y su complemento, $\text{no } A =$ “mayores de 55 años”, y el segundo con otros dos sucesos: $B =$ “ha tenido un ataque al corazón” y su complemento, $\text{no } B =$ “nunca ha tenido un ataque al corazón”.

FIGURA 1: Problema planteado.

Problema. 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 observadas 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

1. ¿Cuántas personas de menor o igual de 55 años nunca han tenido un ataque al corazón?
Explica tu respuesta.
2. ¿Cuántas personas nunca han tenido un ataque al corazón?
Explica tu respuesta.
3. Si elegimos una de estas personas al azar, ¿cuál es la probabilidad de que la persona ha tenido un ataque al corazón?
Explica tu respuesta e indica las operaciones que realizas para responder.
4. Si elegimos una de estas personas al azar, ¿cuál es la probabilidad de ser mayor de 55 años y, al mismo tiempo haber tenido un ataque al corazón?
Explica tu respuesta e indica las operaciones que realizas para responder.
5. Si la persona seleccionada es mayor de 55 años, ¿cuál es la probabilidad de que la persona ha tenido un ataque al corazón?
Explica tu respuesta e indica las operaciones que realizas para responder.

Fuente: Adaptado de Estrada y Díaz (2006).

La primera tarea implica la lectura literal de un dato (nivel 1, leer los datos) representado en la tabla de doble entrada. La segunda hace referencia a la interpretación o integración de los datos, al hacer uso del algoritmo de la suma, para proporcionar un dato que no está representado explícitamente en la tabla (nivel 2, leer dentro de los datos). La tercera está relacionada con el cálculo de la probabilidad de un evento simple (probabilidad simple). La cuarta refiere al cálculo de una probabilidad conjunta. Finalmente, la quinta hace referencia al cálculo de la probabilidad

de que ocurra un evento, dado que ha ocurrido otro (probabilidad condicional). Cabe señalar que este tipo de tareas son similares a las propuestas en el libro de texto de primer año medio, otorgado gratuitamente por el Ministerio de Educación de Chile (GALASSO et al., 2016).

Con la finalidad de evaluar la pertinencia del enunciado del problema y las tareas, se realizó la validación del cuestionario por juicio de expertos y aplicación piloto.

La aplicación del cuestionario se realizó en una sesión de clase, disponiendo de aproximadamente 45 minutos para su resolución. Las profesoras titulares de los cursos (1°, 2°, 3° y 4° medio) colaboraron con la aplicación adoptando un carácter de tipo observador, no involucrándose en los procesos de resolución de las tareas.

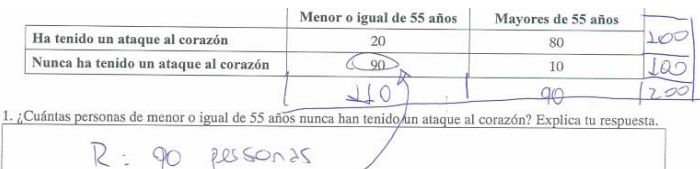
Recogidos los datos, se realizó un proceso cíclico de análisis, con una triangulación de expertos, y se categorizaron las respuestas de cada tarea. Para las respuestas de las tareas 1 y 2, se optó por las categorías: *dominio* y *no dominio*, en caso de que la respuesta sea correcta o incorrecta; además, en la tarea 2 se consideró la categoría *dominio parcial*, cuando la respuesta presenta el dato solicitado sin su justificación explícita, o bien, se proporciona un dato distinto al esperado debido a errores de cálculo en el algoritmo de la suma. Con respecto a las tareas 3, 4 y 5, se optó por las categorías: *correcta*, en la que se agrupan aquellas respuestas que presentan el valor de la probabilidad solicitada, junto con su justificación; *parcialmente correcta*, cuando la respuesta presenta el valor de la probabilidad solicitada, pero sin una justificación explícita, o bien, proporciona una respuesta distinta a la esperada debido a errores de cálculo; e *incorrecta*, cuando la respuesta es errónea y presenta un error, o bien, se proporciona una probabilidad bajo el enfoque subjetivo (BATANERO, 2005). Cabe señalar que en todas las tareas se consideró la categoría *sin respuesta*, cuando no se aborda la tarea y se deja el espacio de la respuesta en blanco.

Análisis y resultados

Tarea 1. Lectura literal de un dato

En la tarea 1, que exige el nivel leer los datos, se esperaba que los estudiantes proporcionaran la frecuencia absoluta doble solicitada, 90 personas. En la Tabla 2 se presentan algunas respuestas, con su clasificación.

TABLA 2: Clasificación de algunas respuestas a la tarea 1.

Categoría	Respuesta / Descripción (D)
Dominio	<p>Estudiante 22, primero medio.</p>  <p>D: Presenta la frecuencia absoluta doble solicitada.</p>
No dominio	<p>Estudiante 18, primero medio.</p> <p>20</p> <p>D: Presenta confusión entre frecuencias absolutas dobles, es decir, confunde $n(A \cap \text{no } B)$ con $n(A \cap B)$.</p>

Fuente: Elaborada por los autores.

A continuación, en la Tabla 3 se resumen los resultados del análisis de las respuestas de los estudiantes a la tarea 1, por categoría y curso.

TABLA 3: Frecuencia (y porcentaje) de dominio del nivel 1, leer los datos, por categoría y curso

Categoría	1°	2°	3°	4°	General
Dominio	20(80)	20(91)	11(78,6)	13(92,9)	64(85,3)
No dominio	3(12)	2(9)	3(21,4)	1(7,1)	9(12)
Sin respuesta	2(8)				2(2,7)
Total	25(100)	22(100)	14(100)	14(100)	75(100)

Fuente: Elaborada por los autores.

La mayor proporción de las respuestas de los estudiantes a la tarea 1 se clasifican en la categoría de *dominio* (80%, 91%, 78,6% y 92,9%, por curso, respectivamente), por lo que esta tarea puede considerarse como lograda por los estudiantes. Por otro lado, estos resultados sugieren que es menos posible que un estudiante de mayor nivel académico no conteste la tarea. Cabe señalar que en todas las respuestas clasificadas en la categoría *no dominio*, los estudiantes confunden $n(A \cap \text{no } B)$ con $n(A \cap B)$, siendo tercero medio aquel curso que presenta mayor porcentaje de respuestas erradas.

Tarea 2. Lectura dentro de los datos

En la tarea 2, que exige el nivel leer dentro de los datos, se esperaba que los estudiantes proporcionaran la frecuencia absoluta marginal solicitada, 100 personas, explicando su respuesta (por ejemplo, aludiendo al algoritmo de la suma “90+10”). En la Tabla 4 se presentan algunas respuestas, con su clasificación.

TABLA 4: Clasificación de algunas respuestas a la tarea 2.

Categoría	Respuesta / Descripción (D)
Dominio	Estudiante 8, segundo medio. <i>100 personas, se sumaron las personas que muoran han tenido um ataque al corazón, 90+10.</i> D: Presenta la frecuencia absoluta marginal solicitada junto con una explicación formal aritmética.
Dominio parcial	Estudiante 6, tercero medio. <i>100 personas</i> D: Presenta la frecuencia absoluta marginal solicitada, pero no explica aritméticamente y/o con enunciados su solución.
No dominio	Estudiante 4, cuarto medio. <i>10 personas según lo explicito</i> D: Presenta confusión entre la frecuencia absoluta marginal solicitada con una frecuencia absoluta doble; confunde $n(\text{no } B)$ con $n(\text{no } A \cap \text{no } B)$.

Fuente: Elaborada por los autores.

En la Tabla 5 se resumen los resultados del análisis de las respuestas de los estudiantes a la tarea 2, por categoría y curso.

TABLA 5: Frecuencia (y porcentaje) de dominio del nivel 2, leer dentro de los datos, por categoría y curso

Categoría	1°	2°	3°	4°	General
Dominio	14(56)	12(54,5)	8(57,1)	9(64,3)	43(57,4)
Dominio parcial	10(40)	10(45,5)	6(42,9)	4(28,6)	30(40)
No dominio				1(7,1)	1(1,3)
Sin respuesta	1(4)				1(1,3)
Total	25(100)	22(100)	14(100)	14(100)	75(100)

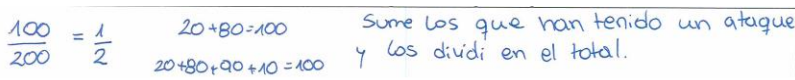
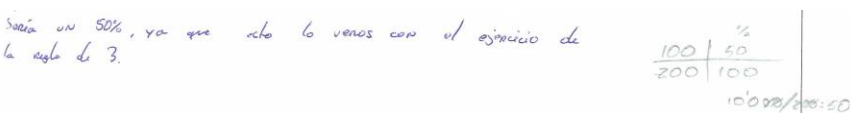
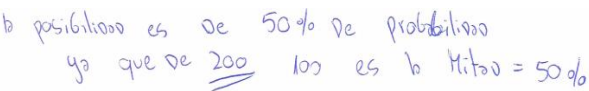
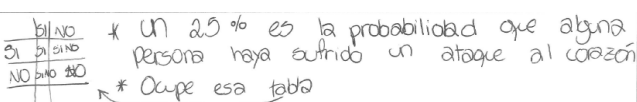
Fuente: Elaborada por los autores.

La mayor proporción de las respuestas de los estudiantes a la tarea 2 se clasifican en la categoría de *dominio* (56%, 54,5%, 57,1% y 64,3% por curso, respectivamente). Además, se evidencia un alto número de respuestas categorizadas como *dominio parcial*, en las que se proporciona el dato solicitado, pero sin presentar explicación aritmética y/o enunciados para validar su respuesta (40%, 45,5%, 42,9% y 28,6% por curso, respectivamente). Lo anterior evidencia que casi la totalidad de los estudiantes presentan un dominio total o parcial del nivel 2 de lectura.

Tarea 3. Cálculo de una probabilidad simple

En la tarea 3 se esperaba que los estudiantes calcularan la probabilidad de un evento simple a partir de la regla de Laplace o el cálculo de la frecuencia relativa marginal por fila, haciendo uso del algoritmo $(a+b)/(a+b+c+d)$: $P(B) = P(\text{una persona haya tenido un ataque al corazón}) = (20+80)/(20+80+90+10) = 100/200 = 0,5$. En la Tabla 6 se presentan algunas respuestas, con su clasificación.

TABLA 6: Clasificación de algunas respuestas a la tarea 3.

Categoría	Respuesta / Descripción (D)
Correcta	<p>Estudiante 13, cuarto medio.</p>  <p>D: El estudiante proporciona la probabilidad correcta, a partir del uso de la regla de Laplace.</p>
	<p>Estudiante 21, tercero medio.</p>  <p>D: El estudiante proporciona la probabilidad correcta, a partir del uso la regla de tres. Calcula la probabilidad a partir de la operación $(100 \cdot 100) / 200$, considerando que el 100% corresponde a 200.</p>
	<p>Estudiante 5, primero medio.</p>  <p>D: El estudiante proporciona la probabilidad correcta, a partir del uso de proporcionalidad de cantidades. Indica que 100 equivale a la mitad del total (200).</p>
Incorrecta	<p>Estudiante 1, cuarto medio.</p>  <p>D: Presenta el sesgo de la equiprobabilidad al asignar la misma probabilidad a cada evento ligado a las frecuencias dobles.</p>

Fuente: Elaborada por los autores.

En la Tabla 7 se resumen los resultados del análisis de las respuestas de los estudiantes a la tarea 3, por categoría y curso.

TABLA 7: Frecuencia (y porcentaje) de respuestas a la tarea 3, por categoría y curso

Categoría		1°	2°	3°	4°	Genera l
Correcta	Regla de Laplace	13(52)	9(40,9)	8(57,1)	8(57,1)	38(50,7)
	Regla de tres	4(16)	10(45,5)	6(42,9)	2(14,3)	22(29,3)
	Proporcionalidad	3(12)			2(14,3)	5(6,7)
Parcialment e correcta	Sin explicación	1(4)	2(9,1)		1(7,1)	4(5,3)
Incorrecta	Equiprobabilida d				1(7,1)	1(1,3)
	Confusión de los casos favorables con los posibles		1(4,5)			1(1,3)
	Confusión P(B) con P(no A ∩ B)	1(4)				1(1,3)
Sin respuesta		3(12)				3(4)
Total		25(100)	22(100)	14(100)	14(100)	75(100)

Fuente: Elaborada por los autores.

La mayoría de las respuestas de los estudiantes a la tarea 3 se clasifican en la categoría *correcta* (80%, 86,4%, 100% y 85,7% por curso, respectivamente). Además, se identifica que: 1) la estrategia más utilizada fue la regla de Laplace en 1°, 3° y 4° medio, y la regla de tres en 2° medio; y 2) fueron pocas las respuestas clasificadas como *parcialmente correctas* e *incorrectas*, presentes en 1°, 2° y 4° medio. Cabe señalar que la presencia del uso de proporcionalidad, aunque mínima, refleja que los estudiantes son capaces de calcular la probabilidad mediante la relación entre diferentes datos de la tabla, considerando el total de la muestra como el 100%.

Tarea 4. Cálculo de una probabilidad conjunta

En la tarea 4 se esperaba que los estudiantes calcularan la probabilidad de la conjunción de dos eventos (probabilidad conjunta), a partir de la regla de Laplace o el cálculo de la frecuencia relativa doble, haciendo uso del algoritmo $(b)/(a+b+c+d)$: $P(\text{no } A \cap B) = P(\text{una persona sea mayor de 55 años y haya tenido un ataque al corazón}) = (80)/(20+80+90+10) = 80/200 = 0,4$. En la tabla 8 se muestran algunas respuestas, con su clasificación.

TABLA 8: Clasificación de algunas respuestas a la tarea 4.

Categoría	Respuesta / Descripción (D)
Incorrecta	<p>Estudiante 17, cuarto medio.</p> <p><i>80% de posibilidad de 100%</i></p> <p>D: Presenta el valor de una probabilidad condicional $[P(\text{no } A B)]$ en lugar de la probabilidad conjunta solicitada $[P(\text{no } A \cap B)]$.</p> <p>Estudiante 19, cuarto medio.</p> <p><i>$\frac{80}{90}$ da quem maior que ^{que} son maiores de 55 han sofrido um ataque do coração. Soma as pessoas de 55 e depois vê o número de pessoas que han sofrido um ataque do coração.</i></p> <p>D: Calcula una probabilidad condicional $[P(B \text{no } A)]$ en lugar de la probabilidad conjunta solicitada $[P(\text{no } A \cap B)]$.</p>

Fuente: Elaborada por los autores.

En seguida, en la tabla 9 se resumen los resultados del análisis de las respuestas de los estudiantes a la tarea 4, por categoría y curso.

TABLA 9: Frecuencia (y porcentaje) de respuestas a la tarea 4, por categoría y curso

Categoría		1°	2°	3°	4°	General
Correcta	Regla de Laplace	11(44)	8(36,4)	8(57,1)	7(50)	34(45,3)
	Regla de tres	3(12)	10(45,5)	5(35,7)	3(21,4)	21(28)
Parcialmente correcta	Sin explicación			1(7,1)		1(1,3)
Incorrecta	Confunde $P(\text{no } A \cap B)$ con $P(\text{no } A B)$				3(21,4)	3(4)
	Confunde $P(\text{no } A \cap B)$ con $P(B \text{no } A)$				1(7,1)	1(1,3)
	Confunde $P(\text{no } A \cap B)$ con $P(A \cap \text{no } B)$	2(8)	1(4,5)			3(4)
	Confusión $P(\text{no } A \cap B)$ con $P(A \cap \text{no } B) + P(\text{no } A \cap B)$	1(4)				1(1,3)
	Enfoque subjetivo	2(8)	1(4,5)			3(4)
Sin respuesta		6(24)	2(9,1)			8(10,7)
Total		25(100)	22(100)	14(100)	14(100)	75(100)

Fuente: Elaborada por los autores.

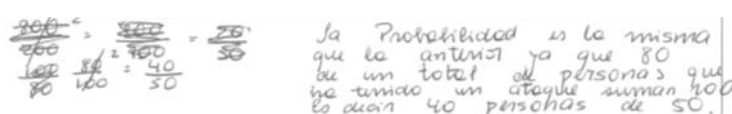
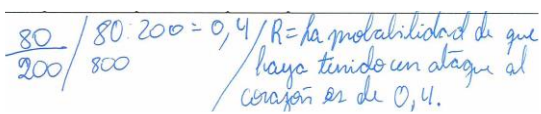
La mayor proporción de las respuestas de los estudiantes a la tarea 4 se clasifican en la categoría *correcta* (56%, 81,9%, 92,8% y 71,4%, por curso, respectivamente), siendo el curso de 3° medio con mayor porcentaje. Las estrategias más usadas fueron la regla de Laplace en 1°, 3° y 4° medio (44%, 57,1% y 50%, respectivamente) y la regla de tres, en 2° medio (45,5%). En relación con las respuestas *parcialmente correctas*, solo se presentó un caso en 3° medio al proporcionar la probabilidad solicitada, pero sin explicación

aritmética y/o enunciados para validar su respuesta; y con respecto a las respuestas *incorrectas*, se presentó la mayor proporción en los estudiantes de 4° medio, al confundir la probabilidad conjunta solicitada con una condicional (28,5%). Cabe destacar que en esta tarea aumenta el número de estudiantes que no la responden con respecto a la anterior.

Tarea 5. Cálculo de una probabilidad condicional

En la tarea 5 se esperaba que los estudiantes calcularan la probabilidad de que ocurra un evento dado que ha ocurrido otro (probabilidad condicional), a partir de la regla de Laplace o el cálculo de la frecuencia relativa condicional por columna haciendo uso del algoritmo $(b)/(b+d)$: $P(B | \text{no } A) = P(\text{una persona haya tenido un ataque al corazón, dado que sea mayor de 55 años}) = (80)/(80+10) = 80/90 = 0,88$. En la tabla 10 se muestran algunas respuestas a esta tarea.

TABLA 10: Clasificación de algunas respuestas a la tarea 5.

Categoría	Respuesta / Descripción (D)
Incorrecta	<p>Estudiante 1, segundo medio.</p>  <p>D: Presenta confusión entre probabilidades condicionales (falacia de la condicional transpuesta), es decir, confunde $P(B \text{no } A)$ con $P(\text{no } A B)$.</p> <p>Estudiante 15, segundo medio.</p>  <p>D: Presenta una probabilidad conjunta $[P(B \cap \text{no } A)]$ en lugar de la condicional solicitada $[P(B \text{no } A)]$.</p>

Fuente: Elaborada por los autores.

En seguida, en la tabla 11 se resumen los resultados del análisis de las respuestas de los estudiantes a la tarea 5, por categoría y curso.

TABLA 9: Frecuencia (y porcentaje) de respuestas a la tarea 5, por categoría y curso

Categoría		1°	2°	3°	4°	General
Correcta	Regla de Laplace	10(40)	4(18,2)	7(50)	9(64,3)	30(40)
	Regla de tres	2(8)	8(36,4)	4(28,6)	2(14,3)	16(21,3)
Parcialmente correcta	Sin explicación			1(7,1)		1(1,3)
Incorrecta	Falacia de la condicional transpuesta				3(21,4)	3(4)
	Confunde $P(B A_c)$ con $P(B \cap A_c)$	2(8)	3(13,6)	2(14,3)	1(7,1)	8(10,7)
	Enfoque subjetivo	3(12)	4(18,2)			7(9,3)
Sin respuesta		6(24)	2(8)	1(4,5)		1(7,1)
Total		25(100)	22(100)	14(100)	14(100)	75(100)

Fuente: Elaborada por los autores.

Se observa que la mayor proporción de las respuestas de los estudiantes a la tarea 5 se clasifican en la categoría *correcta* (48%, 54,6%, 78,6% y 78,6%, por curso, respectivamente), siendo los cursos de 3° y 4° medio con el mayor porcentaje. Así como en las tareas 3 y 4, las estrategias más utilizadas por los estudiantes fueron la regla de Laplace en 1°, 3° y 4° medio (40 %, 50 % y 64,3 %, respectivamente), y la regla de tres en 2° medio (36,4 %). En cuanto a las respuestas *parcialmente correctas*, únicamente se presentó un caso con ausencia de explicación; mientras que, en la categoría *incorrectas*, se presentan con mayor proporción dos errores, la confusión de la probabilidad condicional con una conjunta (12% y 18,2 %, en 1° y 2°, respectivamente) y la confusión de la

dirección de la probabilidad condicional, es decir, confundir $P(B \mid \text{no } A)$ con $P(\text{no } A \mid B)$ (8%, 13,6%, 14,3% y 7,1%, en 1°, 2°, 3° y 4°, respectivamente). Asimismo, se observa un incremento en el número de estudiantes que deja el espacio de la respuesta en blanco en relación con las tareas anteriores.

Conclusiones

Los resultados obtenidos ponen de manifiesto que los estudiantes de educación media presentan en su mayoría *dominio* de los niveles de lectura 1 (leer los datos) y 2 (leer entre los datos) de Curcio, al realizar la lectura literal de la frecuencia doble en la tarea 1 y el procedimiento de la suma para determinar la frecuencia marginal por fila en la tarea 2. Cabe señalar que, en esta última tarea, se presenta un alto porcentaje de respuestas categorizadas como *dominio parcial*, ya que los estudiantes no proporcionan una explicación aritmética y/o enunciados para validar su solución; lo que puede ser causado por no leer la expresión: explica tu respuesta, o bien, por algún aspecto actitudinal del estudiante al no querer justificar su solución. Con respecto a estas tareas, nuestros resultados son similares con los obtenidos en estudios realizados con estudiantes, con profesores en formación y en servicio (DÍAZ-LEVICOY et al., 2016; GARCÍA-GARCÍA et al., 2019; RODRÍGUEZ; SANDOVAL, 2012) y, además, sugieren la consideración de la justificación, en lo que respecta tarea relacionada con el nivel 2 de Curcio, como una innecesaria (40% de respuestas sin justificación) a pesar de ser una componente necesaria de la alfabetización estadística.

En relación con las tareas que involucran cálculo de probabilidades, observamos que la mayoría de los estudiantes responden correctamente haciendo uso de la regla de Laplace como estrategia más utilizada en los cursos 1°, 3° y 4° medio, y de la regla de tres en 2° medio; este alto porcentaje de respuestas *correctas* coincide con los resultados reportados en Contreras et al. (2010) y Estrada y Díaz (2006; 2007). Los errores con mayor proporción de ocurrencia en las respuestas de los estudiantes fueron confundir la probabilidad

conjunta solicitada con una probabilidad condicional o con otra conjunta (tarea 4) y confundir la probabilidad condicional solicitada [$P(B \mid \text{no } A)$] con su inversa [$P(\text{no } A \mid B)$], o con una conjunta (tarea 5); estos coinciden con los reportados por Estrada y Díaz (2006; 2007) y Contreras et al. (2010). Por otro lado, en esta investigación se presentan respuestas bajo el enfoque subjetivo de la probabilidad, esto posiblemente por el nivel educativo de los estudiantes, al otorgar un valor a la probabilidad de acuerdo con su grado de creencia personal (BATANERO, 2005) y el cálculo de probabilidades utilizando proporcionalidad, lo que sugiere que los participantes pueden calcular las probabilidades por métodos alternativos a la regla de Laplace o el uso de regla de tres, que tienen una naturaleza más algorítmica. En cuanto al porcentaje de estudiantes que no responden a las tareas de cálculo de probabilidades, y de aquellos que proporcionan la probabilidad solicitada sin explicación, estos son menores con respecto a los reportados en estudios similares (CONTRERAS et al., 2010; ESTRADA; DÍAZ, 2006; 2007); además los estudiantes de primero medio son los únicos que no responden las tareas de menor complejidad (1 a 3), lo cual se puede atribuir a que aún no han desarrollado lo suficiente las competencias y habilidades para la búsqueda de soluciones a tareas como las presentadas.

Lo reportado en esta investigación, acerca del dominio de los primeros niveles de lectura, así como de las estrategias y errores en el cálculo de probabilidades, dan la base para el diseño de experiencias de aprendizaje o secuencias didácticas donde se pongan de manifiesto los diversos conceptos y procesos matemáticos en torno a las tablas de doble entrada y, con ello, lograr una mayor comprensión por el estudiante; por ejemplo, abordar los diversos significados de la probabilidad, hacer hincapié del uso de los enfoques clásico y frecuencial para calcular o estimar probabilidades a partir de datos recopilados en su entorno, entre otros aspectos. Por otro lado, el número de estudiantes que no responden las tareas deja entrever la necesidad de indagar sus causas, como posible investigación a futuro. En general, nuestros resultados manifiestan la necesidad de reforzar la formación del estudiantado de educación media (desde 1° hasta 4° medio)

preferentemente en el cálculo de probabilidades, fomentando el desarrollo de componentes esenciales para la alfabetización probabilística y estadística.

Referencias

- BATANERO, C. Significados de la probabilidad en la educación secundaria. *Relime. Revista Latinoamericana de Investigación en Matemática Educativa*, v.8, n.3, p.247-263. 2005.
- CONTRERAS, J. M. et al. Dificultades de futuros profesores en la lectura y cálculo de probabilidades en tablas de doble entrada. En MORENO, M. et al. (Eds.), *Investigación en Educación Matemática*, v.XIV, n.1, p.271-280. 2010.
- CURCIO, F. R. *Developing graph comprehension*. Reston, VA: National Council of Teachers of Mathematics [NCTM]. 1989.
- DÍAZ, C.; DE LA FUENTE, E. Conflictos semióticos en el cálculo de probabilidades a partir de tablas de doble entrada. *Biaix*, v.24, n.1, p. 85-91. 2005.
- DÍAZ-LEVICOY, D. et al. Lectura de tablas estadísticas por futuras maestras de educación infantil. *Educação Matemática Pesquisa*, v.18, n.3, p.1099-1115. 2016.
- ESTRADA, A.; DÍAZ, C. 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. Proceedings...* Salvador de Bahia, Brasil. 2006.
- ESTRADA, A.; DÍAZ, C. Errores en el cálculo de probabilidades en tablas de doble entrada en profesores en formación. *Uno. Revista de Didáctica de las Matemáticas*, n.44, p.48-57. 2007.
- FRIEL, S.; CURCIO, F.; BRIGHT, G. Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, v.32, n.2, p.124-158, 2001.
- GAL, I. Adult's statistical literacy: Meaning, components, responsibilities. *International Statistical Review*, v.70, n.1, p.1-25. 2002.
- GAL, I. Towards "probability literacy" for all citizens: building blocks and instructional dilemmas. En JONES, G. (Ed.) *Exploring probability in school: Challenges for Teaching and Learning*. Boston, MA.: Springer, 2005, p. 39-63.
- GAL, I. Understanding statistical literacy: About knowledge of contexts and models. En J. M. CONTRERAS; M. M. GEA; M. M. LÓPEZ-MARTÍN; E. MOLINA-PORTILLO (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. Granada, España. 2019.

GALASSO, B., MALDONADO, L., Y MARAMBIO, V. Texto del estudiante. Matemática. Primero Medio. Santiago: Santillana. 2016.

GARCÍA-GARCÍA, J. I. et al. Comprensión de una tabla estadística por estudiantes universitarios en México y Chile. *Revemat. Revista Eletrônica de Educação Matemática*, v.14, p.1-16. 2019.

GEA, M. M. et al. Construcción y lectura de la tabla de doble entrada por profesores de Educación Primaria en formación. *Educação Matemática Pesquisa: Revista do Programa de Estudos Pós-Graduados em Educação Matemática*, v.22, n.1, p.348-370. 2020.

HERNÁNDEZ, R.; FERNÁNDEZ, C.; BAPTISTA, P. Metodología de la Investigación. México: McGraw Hill. 2010.

KRIPPENDORFF, K. Metodología de análisis de contenido. Teoría y práctica. Barcelona: Paidós Comunicación. 1997.

MINEDUC [MINISTERIO DE EDUCACIÓN DE CHILE]. Matemática. Programa de Estudio. Primero Medio. Santiago: MINEDUC. 2016.

POGGIOLI, L. Estrategias de resolución de problemas. Serie enseñando a aprender. Caracas: Fundación Polar. 1999.

RICO, L. Errores y dificultades en el aprendizaje de las Matemáticas. En KILPATRIK, J.; GÓMEZ, P.; RICO, L. (Eds.). Educación Matemática. Errores y dificultades de los estudiantes. Resolución de problemas. Evaluación. Historia. Bogotá: Grupo Editorial Iberoamérica, 1996. p.69-108.

RODRÍGUEZ, F; SANDOVAL, P. Habilidades de codificación y descodificación de tablas y gráficos estadísticos: un estudio comparativo en profesores y alumnos de pedagogía en enseñanza básica. *Avaliação: Revista de la avaliação da Educação Superior*, v.17, n.1, p. 207-235. 2012.

RODRIGUEZ-ALVEAL, F.; DIAZ-LEVICOY, D.; VASQUEZ, C. Evaluación de la alfabetización probabilística del profesorado en formación y en activo. *Estudios pedagógicos*, v.44, n.1, p. 135-156. 2018.

SÁNCHEZ, E. La probabilidad en el programa de estudio de matemáticas de la secundaria en México. *Educación matemática*, v.21, n.2, p.39-77. 2009.

SCHILD, M. Statistical literacy survey analysis: Reading graphs and tables of rates and percentages. En *Sixth International Conference on Teaching Statistics. Proceedings...* 2006.

SOCAS, M. Dificultades, obstáculos y errores en el aprendizaje de las Matemáticas en la Educación Secundaria. En RICO, L. (Ed.). *La educación matemática en la enseñanza secundaria*. Barcelona: Horsori, 1997. p.125-144.

Recibido en agosto de 2023.

Aprobado en octubre de 2023.