

# DIFICULTADES DE LOS ESTUDIANTES PARA LA ARTICULACIÓN DE LOS DISTINTOS SIGNIFICADOS DE VARIABLE EN UN CURSO BÁSICO DE ESTADÍSTICA UNIVERSITARIA

## STUDENTS' DIFFICULTIES FOR CONNECTING THE DIFFERENT MEANINGS OF VARIABLE IN A BASIC COURSE OF UNIVERSITY STATISTICS

**Alvaro Cortínez, Armando Albert, Blanca Ruiz**

Universidad de Tarapacá (Chile). Tecnológico de Monterrey (México)  
acortinez@academicos.uta.cl, albert@tec.mx, bruiz@tec.mx

### Resumen

La Estadística ocupa un lugar cada vez más importante en una sociedad que aumenta continuamente su producción de información cuantitativa. Sin embargo, su enseñanza presenta dificultades para alcanzar sus objetivos. Una de esas dificultades tiene que ver con un concepto clave para el planteamiento de modelos de solución a problemas: la variable. Los estudiantes se encuentran desde los primeros niveles de enseñanza hasta el universitario con el concepto de variable. Al llegar a un curso de Estadística, la idea de variable toma nuevos significados vinculados a los datos y en probabilidad a una función. Este trabajo muestra un estudio exploratorio en estudiantes de ingeniería y medicina realizado a lo largo de su primer curso de estadística sobre sus distintas interpretaciones que le dan a la variable. Finalmente, se muestra que un concepto como el de variable, casi obvio en el discurso escolar, presenta, en realidad, una gran complejidad epistémica, didáctica y cognitiva que demanda ser atendida.

**Palabras clave:** enseñanza de la probabilidad y estadística, variable

### Abstract

Statistics occupies a more and more relevant place in a society that continuously increases its production of quantitative information. However, its teaching presents difficulties to achieve its objectives. One of those difficulties has to do with a key concept for the approach of problem-solving models: the variable. The students meet the concept of variable from the first levels of education up to the university. When facing a Statistics course, the idea of variable takes new meanings linked to the data, and in probability, to a function. This paper shows an exploratory study on the different interpretations engineering and medical students give to the variable during their first statistics course. Finally, it is shown that a concept like that of variable, almost obvious in school discourse, actually presents, a great epistemic, didactic, and cognitive complexity that demands special attention.

**Key words:** probability and statistics teaching, variable

## ■ Introducción

Desde los primeros años de la enseñanza básica hasta el universitario, los estudiantes se encuentran con el concepto de *variable*, pero sus significados cambian según la disciplina abordada tal como el álgebra, el cálculo, la probabilidad o la estadística. Sin embargo, deben saber articular y diferenciar los distintos significados en el proceso de plantear un modelo de solución al problema por resolver. Por ejemplo, la *variable* en álgebra es un elemento que no tiene un valor fijo y que se vincula con otros elementos a través de relaciones de equivalencia. En cambio, en cálculo la *variable* está sujeta a relaciones de dependencia-independencia, así como a procesos de convergencia. En estadística se habla de variable estadística -vinculada a los datos- con su clasificación (cualitativa-cuantitativa, y sus escalas) y en probabilidad de variable aleatoria, muy distante de la variable algebraica o de cálculo, pues en realidad se trata de una función que asocia elementos del espacio muestral con números reales. Existen algunos estudios que abordan el problema didáctico de la variable tal como el de Ursini, Trigueros y Lozano (2000) donde, desde la perspectiva del álgebra, la variable ya presenta complejidad por sus distintos significados tales como ser una incógnita, un número general y ocupar un papel central en la relación funcional de dependencia – independencia. Sin embargo, según estos autores, los estudiantes de enseñanza media, entre 12 y 19 años, no adquieren la capacidad de interpretar, simbolizar y manipularla adecuadamente. Estudios más recientes como el de Herrera, Cuesta y Escalante (2016) obtienen resultados similares, añadiendo la dificultad de los estudiantes de no lograr asociar o aplicar el conocimiento que se tiene a las tareas propuestas debido al escaso nivel de razonamiento que no le permite procesar las relaciones entre conceptos. Por otra parte, Ruiz y Albert (2013) muestran una desvinculación en el discurso del aula y los libros de texto entre la variable estadística y la variable aleatoria, pero su tratamiento didáctico podría ayudar a los estudiantes universitarios a superar algunas dificultades de modelación estocástica a problemas vinculados con la realidad. Toda esta complejidad se ve reflejada en los cursos introductorios de Estadística del nivel universitario y, por esta razón se hace necesarios estudios que la aborden para esclarecerla. En este sentido, esta investigación busca aportar elementos descriptivos que profundicen el conocimiento del problema desde la perspectiva del estudiante con relación a su concepción de variable a lo largo de un curso introductorio de Estadística universitaria.

## ■ Marco teórico

El estudio se realiza desde la perspectiva del *pensamiento estadístico* iniciado por Mallows (1998) que vincula el pensamiento estadístico con la relación entre los datos y un problema de contexto real. Wild y Pfannkuch (1999) ubican el pensamiento estadístico en todo el proceso investigación desde la planeación para la toma de datos, pasando por un modelo estadístico o probabilístico, hasta la interpretación, de forma semejante a cómo lo haría un profesional de la estadística. Desde entonces esta perspectiva ha sido retomada por diferentes autores tales como Behar (2018) quien resalta la importancia del contexto para un pensamiento estadístico crítico. Por otra parte, el objeto didáctico de estudio es analizado desde una perspectiva multidimensional: *epistemológica*, de *enseñanza* y *cognitiva*.

### *Dimensión epistemológica*

La variable aleatoria como tal, se comenzó a usar informalmente en el siglo XVII, en trabajos de Jacob Bernoulli, Abraham de Moivre y Pierre Laplace. Fue Poisson, en 1937, en su trabajo *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile*, quien usó por primera vez la expresión variable aleatoria como una función. Posteriormente, Chebyshev (1846 y 1887) utilizó de manera más explícita las ideas de variable aleatoria y de función de densidad, nombrando a las variables aleatorias como cantidades y magnitudes. Liapunov (1900 y 1901) identificó las variables aleatorias y las enunció, trabajando con la función de probabilidad prácticamente como se hace hoy en día.

Formalmente, la variable aleatoria se puede definir de la siguiente forma (Cuadras 1999, citado por Ruiz y Albert, 2013):

Sea  $(\Omega, \mathcal{A}, \mathcal{P})$  un espacio de probabilidad y  $R$  el cuerpo (o campo) de los números reales. Se dice que la aplicación:

$$\begin{aligned} \xi: \Omega &\rightarrow R \\ \omega &\rightarrow \xi(\omega) \in R \end{aligned}$$

que a cada suceso elemental hace corresponder un número real, es una variable aleatoria si para todo número real  $x$ , se verifica la relación:

$$A = \{\omega | \xi(\omega) \leq x\} \in \mathcal{A}$$

es decir, se verifica que  $A$  es un suceso.

Durante mucho tiempo la variable aleatoria fue vista como el resultado numérico de un experimento aleatorio ( $\xi(\omega) = \omega$ ) y llamada simplemente variable y no vista como lo que es actualmente, una función. Esta es una dificultad epistemológica que también se presenta en los estudiantes (Ruiz, 2013).

### *Dimensión de enseñanza*

El concepto de variable es abordado desde los inicios del sistema educativo y no siempre con el mismo éxito. Existen diversos estudios acerca de la variable y su comprensión, especialmente en las áreas de álgebra. La mayoría de ellos se basan en el Modelo 3uv (Ursini y Trigueros, 2006). Estas autoras consideran que la solución competente de los problemas algebraicos requiere un manejo flexible de tres usos de la variable: la incógnita específica, número general y relación funcional. Escalante y Cuesta (2012) realizan una experiencia con estudiantes universitarios, centrados en la comprensión de conceptos algebraicos y de variable en particular, llegando a la conclusión que la incomprensión del concepto, se mantiene igual desde la secundaria hasta la universidad; "los estudiantes continúan evitando cualquier acercamiento algebraico y retornan a procedimientos de carácter aritmético, en especial cuando el problema implica cierto nivel de razonamiento de las situaciones y los diferentes contextos en que se presenta la tarea" (p. 128). Por otra parte, cuando los estudiantes llegan a su primer curso de estadística universitaria se enfrentan no sólo con una variedad de significados no siempre triviales como variable de interés, variable de observación, variable discreta, variable continua, variable según la escala nominal, ordinal, de intervalo o de razón, sino con la complejidad de la variable estadística, variable aleatoria y el estadístico como una variable aleatoria muestral. Incluso, la definición de variable aleatoria es enseñada de diferentes formas: "Una variable aleatoria es una variable que toma valores numéricos determinados por el resultado de un experimento aleatorio" (Newbold, Carlson y Thorne, 2008, p. 146). "Cuando los valores obtenidos surgen como un resultado de factores aleatorios, de modo que no pueden ser predichos exactamente por adelantado, la variable se llama variable aleatoria" (Daniel y Cross, 2018, p.4). De lo anterior puede verse la necesidad de que la variable sea estudiada formalmente por la didáctica para construir propuestas más acordes con su complejidad. En muchos lugares de Latinoamérica, a la variable vinculada con el recuento de los datos, se le nombra "variable" sin el calificativo de "estadística". Esto ocasiona que en la enseñanza no se haga explícita la diferencia entre la variable aleatoria y estadística y por lo tanto, se confundan (Ruiz, 2013).

### *Dimensión cognitiva*

Diversos estudios se han hecho sobre las dificultades que tienen los estudiantes sobre el concepto de variable en general. Tal es el caso de Escalante y Cuesta (2012). Su estudio muestra que, efectivamente, existen dificultades para realizar una lectura analítica de los enunciados verbales y serios obstáculos en el proceso de traducción de los lenguajes natural, aritmético y geométrico al lenguaje algebraico. Por su parte, también hay algunos reportes sobre las dificultades alrededor de la variable aleatoria como el problema de la transferencia de la propiedad de aleatoriedad a la variable dependiente a partir de variables independientes aleatorias (Ruiz, Albert y Batanero, 2006) y el tránsito entre la variable estadística que describe una distribución de los datos y la variable aleatoria como modelo de distribución de probabilidad (Ruiz y Albert, 2013).

## Metodología

Con el propósito de ahondar en la dimensión cognitiva, este estudio busca explorar las nociones de variable en estudiantes de los primeros semestres de su carrera universitaria que están en su primer curso de Estadística. Esto se hará en dos grupos de estudiantes de diferente perfil: del área administrativa y de ciencias de la salud.

### El contexto del estudio

El estudio se aplicó a dos grupos de estudiantes: Ingeniería Comercial y Medicina en una universidad del norte de Chile. Ambos grupos de estudiantes están cursando el segundo año de la carrera y tienen la asignatura de Estadística (Probabilidad y Estadística los primeros y Bioestadística los segundos). En el curso de Probabilidad y Estadística hay este año 45 estudiantes y en el de Bioestadística, 32 estudiantes. De ellos, 42 y 24 estudiantes, respectivamente, respondieron el cuestionario.

### El instrumento

Se implementó un cuestionario (Figura 1) diseñado en dos partes: la primera, referida a sus nociones de variable desde una distribución de probabilidad familiar para ellos; la segunda, que a su vez consta de dos partes: indaga las nociones desde una perspectiva contextual y gráfica. Todos los estudiantes contaron con media hora para resolverlo de manera individual. El cuestionario se aplicó al final del semestre académico en donde se trataron los temas de variable estadística y aleatoria.

<p><b>PRIMERA PARTE: CUESTIONES</b></p> <ol style="list-style-type: none"> <li>¿Cuándo recuerda usted haber visto por primera vez el concepto de variable?</li> <li>Dada la siguiente función de probabilidad, indique cuál es la variable y cuál es el parámetro:  <math display="block">f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}</math> </li> <li>"Es un valor fijo y desconocido que caracteriza al modelo de probabilidad". Esta es una versión de la definición de:             <ol style="list-style-type: none"> <li>Dar un ejemplo de una variable cuantitativa discreta infinita.</li> </ol> </li> <li>Dar un ejemplo de una variable cuantitativa continua.</li> <li>¿Es lo mismo una variable matemática (por ejemplo en la ecuación de la recta) y una variable estadística (por ejemplo en la función de densidad)? Explique.</li> </ol> <p><b>SEGUNDA PARTE: PROBLEMAS</b></p> <ol style="list-style-type: none"> <li>La Cámara de Comercio de la Región de Arica y Parinacota, por estudios hechos anteriormente, sabe que las mujeres adultas poseen en promedio 18 pares de zapatos. Una zapatería, Los tres García, duda de este resultado pues cree que en la actualidad el promedio es menor a 18 y se da a la tarea de verificarlo. Para eso hace un muestreo aleatorio con 50 mujeres adultas y se le pregunta cuántos pares de zapatos tienen. De la muestra resultó el promedio de 15 pares de zapatos con una desviación estándar de 5 pares.              ¿Cuál es la variable de observación? Señale con una X la respuesta correcta en el paréntesis de la izquierda:              ( ) Diseñar una estrategia exitosa.              ( ) El excesivo número de zapatos que las mujeres adultas de Tarapacá tienen.              ( ) Saber si el promedio actual de pares de zapatos por mujer adulta de Tarapacá es menor que en el pasado.              ( ) El número de pares de zapatos por mujer adulta de Arica y Parinacota              ( ) Otra: _____           </li> <li>Se estudia si se amplía un puente de la Ruta 5 Norte pero el Seremi de Obras públicas sostiene que no es necesario porque en promedio el flujo vehicular ha sido siempre de 30 autos por minuto en hora punta, lo que está dentro de la capacidad del puente. Un ingeniero está en desacuerdo con esta información y piensa que en realidad actualmente es mayor. Para verificarlo, él decide hacer un estudio de flujo vehicular y encuentra que en 30 mediciones resultó un promedio de 40 autos por minuto en hora punta. Hallar:             <ol style="list-style-type: none"> <li>¿Cuál es la variable de interés? _____</li> <li>¿Cuál sería la población relativa a de esta variable? _____</li> <li>Indique el nombre del parámetro de interés en este caso _____ y su valor _____</li> <li>Indique el nombre del estadístico de interés en este caso _____ y su valor _____</li> </ol> </li> <li>Un fabricante de pasta dental quiere saber cuánto están dispuestos a pagar adicionalmente sus clientes por su nuevo dentífrico. Su anterior pasta dental tenía un costo promedio de 2050 pesos. Para eso hace una encuesta con 20 personas elegidas al azar de entre sus clientes en varios supermercados del entorno y encuentra los siguientes resultados:              +180, +280, +70, +210, +280, +70, +35, +180, +35, +180,              +105, +350, +35, +105, +140, +180, +70, +70, +105, +35.              a) ¿Cuál es el problema en este caso?              b) ¿Cuál es la población?           </li> </ol>	<ol style="list-style-type: none"> <li>¿Cuál es la variable de observación?</li> <li>Identifique un dato de la variable de observación</li> <li>¿Qué tipo de variable es? (cualitativa/cuantitativa discreta cuantitativa continua)</li> <li>¿A qué escala pertenecen los datos y por qué? (Nominal/Ordinal /Intervalo/Razón)</li> <li>¿Cuál información dada en el problema corresponde a un parámetro? Indique su nombre y su valor.</li> </ol> <p>4. El histograma de los datos del problema anterior es el siguiente:</p> <div style="text-align: center;"> </div> <ol style="list-style-type: none"> <li>¿Cuál es la variable de observación según el histograma?</li> <li>Si se sabe que la frecuencia es de 4, a qué valor del incremento en pesos corresponde?</li> <li>¿Cuántas variables estadísticas están representadas en el gráfico?</li> </ol> <p>5. Si la curva suave representa ahora a una función aproximada a la distribución de los datos, cuya área total bajo ella es 1, en este caso, ¿habría alguna diferencia entre la variable horizontal del histograma y la variable horizontal de la función? Explique.</p> <div style="text-align: center;"> </div>
---	--

Figura 1. Cuestionario implementado a estudiantes de Ingeniería comercial y Medicina.

## ■ Resultados

### Primera Parte: Nociones generales

La pregunta 1 fue *¿Cuándo recuerda usted haber visto por primera vez el concepto de variable?* Los resultados se muestran en la *Tabla 1* siguiente:

**Tabla 1.** *Sobre cuándo recuerdan los estudiantes haber visto el concepto de variable*

Nivel	Ingeniería Comercial	Medicina
Enseñanza Básica	2 (4.8%)	3 (12.5%)
Enseñanza Media	25 (59.5%)	17 (70.8%)
Preparación Prueba de Selección Universitaria	2 (4.8%)	2 (8.3%)
Primer año Universidad	2 (4.8%)	0
Segundo año Universidad	11 (26.2%)	2 (8.3%)
Total	42	24

El concepto de *variable* está presente desde que los estudiantes se encuentran con álgebra en sus asignaturas de matemática de Enseñanza Básica. Sin embargo, en el cuestionario, los estudiantes la asocian más frecuentemente a Enseñanza Media (bachillerato). Y, más aún, algunos tienen la percepción de encontrarse por primera vez con este concepto en Segundo Año de la Universidad, especialmente los estudiantes de Ingeniería Comercial. Cabe señalar que ellos ya tuvieron 2 asignaturas de matemática en el nivel universitario, donde han conocido los elementos de álgebra y, en particular, de función.

Respecto a la *pregunta 2: Dada la siguiente función de probabilidad, indique cuál es la variable y cuál es el parámetro:*

$$f(x) = P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

Los resultados se muestran en la *Tabla 2* siguiente:

**Tabla 2.** *Sobre identificar de una función de probabilidad la variable y los parámetros*

Resultado	Ingeniería Comercial	Medicina
Respuesta Incorrecta	11 (26.2%)	2 (8.3%)
Acierta un elemento	5 (11.9%)	5 (20.8%)
Acierta dos elementos	11 (26.2%)	16 (66.7%)
Acierta todos los elementos	12 (28.6%)	0
No responde	3 (7.1%)	1 (4.2%)
Total	42	24

Como puede observarse, ningún estudiante de Medicina acertó a los tres elementos. Por su parte, hay un 26.2% de estudiantes de Ingeniería Comercial que respondieron de forma incorrecta. Entre las respuestas, muchos consideraron a  $f(x) = P(X = x)$  como la variable y a la función misma como el parámetro. En ambos grupos hubo

personas que tienen una clara confusión acerca del papel que juega  $n$ , el tamaño de la muestra, considerándolo en algunos casos como una variable y en otros no identificándolo. En Figueroa y Aznar (2017) se propone situar a los estudiantes en problemas con contexto a fin de mejorar su apreciación sobre la variable, idea que se retoma en preguntas posteriores de este análisis.

*Pregunta 3: Es un valor fijo y desconocido que caracteriza al modelo de probabilidad.* Los resultados se muestran en la Tabla 3.

**Tabla 3.** Sobre qué es un valor fijo y desconocido que caracteriza al modelo de probabilidad

Resultado	Ingeniería Comercial	Medicina
Respuesta Incorrecta	25 (59.5%)	17 (70.8%)
Respuesta Correcta	2 (4.8%)	7 (29.2%)
No responde	15 (35.7%)	0
Total	42	24

El concepto de *parámetro* presenta aún más problemas que el de variable, esto concuerda con los resultados de Cortinez, Alamilla, Albert y Ríos (2015). En Ingeniería Comercial, sólo dos estudiantes identifican la definición clásica de parámetro, mientras que, en Medicina, son siete los que conocen esta definición. Entre las respuestas, se encuentran "variable aleatoria", "variable", "Poisson", "variable independiente", "función de densidad".

De las *preguntas 4 y 5* sobre dar un ejemplo de una variable cuantitativa discreta infinita y dar un ejemplo de una variable cuantitativa continua, los resultados se muestran en la Tabla 4.

**Tabla 4.** Sobre dar ejemplos de variable cuantitativa discreta infinita y continua

Resultado	Variable Cuantitativa discreta infinita		Variable cuantitativa continua	
	Ingeniería Comercial	Medicina	Ingeniería Comercial	Medicina
Incorrecto	12 (28.6%)	5 (20.8%)	9 (21.4%)	3 (12.5%)
Está la idea	5 (11.9%)	0	6 (14.3%)	6 (25.0%)
Es correcto	19 (45.2%)	19 (79.2%)	23 (54.8%)	15 (62.5%)
No responde	6 (14.3%)	0	4 (9.5%)	0
Total	42	24	42	42

Estudios previos muestran la complejidad conceptual y cognitiva de la variable como el caso del trabajo de Ruiz (2013) y, en este caso, aunque los estudiantes de Medicina presentan menos problemas a la hora de ejemplificar tipos de variables, se presentaron errores comunes en ambos grupos, no reportados antes, tales como el de dar como ejemplo de variable daban valores. Por ejemplo: "En el curso hay 50 alumnos; 30 son mujeres y 20 son hombres". Esto muestra, en efecto, la complejidad de la variable en su proceso de aprendizaje.

De la *pregunta 6*, ¿Es lo mismo una variable matemática (por ejemplo, en la ecuación de la recta) y una variable estadística (por ejemplo, en la función de densidad)? Explique, se obtuvieron los siguientes resultados expresados en la Tabla 5.

**Tabla 5.** *Sobre si es lo mismo una variable matemática que una estadística*

Resultado	Ingeniería Comercial	Medicina
No	20 (47.6%)	22 (91.7%)
Sí	6 (14.3%)	1 (4.2%)
No responde	16 (38.1%)	1 (4.2%)
Total	42	24

Resultan muy interesantes las diversas respuestas que entregan los estudiantes. En general, tienden a considerar la variable matemática como fija, en sus palabras: “un valor exacto”, mientras que la estadística la clasifican como “aleatoria o inexacta”.

*Segunda parte: Adentrándose en el contexto*

En la tabla 6 se muestran los resultados obtenidos en el *problema 1*. Se hace un muestreo aleatorio con 50 mujeres adultas y se les pregunta cuántos pares de zapatos tienen. ¿Cuál es la variable de observación?

**Tabla 6.** *La variable de observación en el contexto de un problema*

Resultado	Ingeniería Comercial	Medicina
Diseñar una estrategia exitosa.	0	0
El excesivo número de zapatos que las mujeres adultas de Tarapacá tienen.	1 (2.4%)	0
Saber si el promedio actual de pares de zapatos por mujer adulta de Tarapacá es menor que en el pasado.	5 (11.9%)	5 (20.8%)
El número de pares de zapatos por mujer adulta de Arica y Parinacota	26 (61.9%)	18 (75.0%)
Otra	8 (19.0%)	1 (4.2%)
No responde	2 (4.8%)	0
Total	42	24

Algunas de las respuestas que los estudiantes dieron como "Otra", coinciden con las alternativas dadas. Se puede observar en la Tabla 6 que la mayoría de los estudiantes aciertan, esto es que reconocen como la variable de observación al número de pares de zapatos por mujer adulta de Arica y Parinacota. Se pueden observar mejores resultados para identificar la variable en este problema de contexto que en la pregunta 2 de la primera parte, sin contexto, tal como lo reportan Figueroa y Aznar (2017).

Sobre el *problema 2* de la *segunda parte* del cuestionario, relativa a la ampliación de puente en Ruta 5 Norte los resultados fueron los siguientes:

En la identificación de la *variable de interés*, el resultado fue que todos los estudiantes de Medicina, y todos menos uno de Ingeniería Comercial, muestran que entendieron el problema e identificaron la variable de interés, con expresiones como "flujo vehicular", "número de vehículos", "número medio de vehículos". Sin embargo, al identificar cuál era la *población*, tuvieron serias dificultades para identificarla. En Ingeniería Comercial, menos de 10 alumnos dieron una respuesta más o menos adecuada. El resto se limitó a dar un número, o volver a mencionar

la variable. Casi la mitad de los estudiantes de Medicina entregó una respuesta aceptable, los demás cayeron en los mismos errores que el otro grupo.

Sobre identificar el nombre y valor del parámetro y el estadístico, los resultados se muestran en la siguiente Tabla 7.

*Tabla 7. Sobre la identificación del nombre y valor del parámetro y el estadístico*

Resultados	Nombre y valor del parámetro		Nombre y valor del estadístico	
	Ingeniería Comercial	Medicina	Ingeniería Comercial	Medicina
Incorrecto	12 (28.6%)	2 (8.3%)	0	0
Acierta sólo el concepto	0	0	0	0
Acierta sólo el valor	14 (33.3%)	8 (33.3%)	5 (14.3%)	4 (16.7%)
Acierta ambos: concepto y valor	0	14 (58.3%)	4 (9.5%)	20 (83.3%)
No responde	16 (38.1%)	0	32 (76.2%)	0
Total	42	24	42	42

En Albert y Ruiz (2014) se muestra que los estudiantes tienen dificultades para distinguir al estadístico como variable aleatoria. Una de esas dificultades, como lo muestra este apartado, consiste en confundir el parámetro con la variable aleatoria, aunque los estudiantes de medicina, que ya habían visto inferencia estadística antes de responder a este cuestionario, obtuvieron mejores resultados. La mayoría de los estudiantes de Ingeniería Comercial contestaron de forma incorrecta o simplemente no contestaron.

Con relación al *problema 3* de la *segunda parte*, los resultados se muestran en la Tabla 8.

*Tabla 8. Sobre un contexto confirmatorio acerca de los conceptos básicos*

Resultados correctos sobre:	Ingeniería Comercial	Medicina
El Problema	17 (40.5%)	21 (87.5%)
Población	15 (35.7%)	20 (83.3%)
Variable de observación	25 (59.5%)	15 (62.5%)
Identificar un dato	7 (16.7%)	11 (45.8%)
Tipo de variable	34 (81.0%)	22 (91.7%)
Escala	2 (4.8%)	13 (54.2%)
Identificación del parámetro	17 (40.5%)	18 (75%)
Total	42	24

De manera similar a los resultados anteriores, se pudo confirmar que los estudiantes de Medicina tienen una mejor percepción de lo que son los conceptos básicos, presentando ambos grupos mayores dificultades con la especificación de la Escala. Aún así, con respecto a la identificación de la variable, sigue habiendo un porcentaje importante de errores de alrededor del 40% en cada grupo.



Los estudiantes de Ingeniería tienden a confundir el concepto con su valor. Por ejemplo, no pocos dieron un tamaño poblacional (generalmente 20, erróneamente) cuando se les pidió identificar a la Población. También tuvieron problemas con la identificación de un dato.

### *Interpretación en contexto gráfico*

Respecto al *problema 4 de la segunda parte*, dado un histograma, se muestran los resultados en la Tabla 9.

**Tabla 9.** *Sobre la interpretación de un histograma*

Resultados correctos sobre:	Ingeniería Comercial	Medicina
Variable de observación	33 (78.6%)	22 (91.7%)
El valor del incremento	39 (92.9%)	23 (95.8%)
Variabes estadísticas	20 (47.6%)	16 (66.7%)
Total	42	24

En general los estudiantes supieron identificar correctamente la variable de observación, así como el valor de incremento para la frecuencia 4. Esto indica que la identificación del histograma ha sido adecuada. Se presenta, sin embargo, un problema en la enumeración de las variables. Alrededor de la mitad de los estudiantes declararon que había 2 variables estadísticas, considerando la frecuencia como una de ellas. Es posible que sea por la influencia de su matemática previa, sin embargo, en el caso de estadística, se trata una sola variable estadística y la frecuencia es sólo una variable auxiliar de agrupación. Pero también, hubo varios estudiantes, especialmente de Ingeniería Comercial, que confundieron el número de variables con el número de intervalos del histograma.

Y, finalmente, del *problema 5 de la segunda parte*, donde se le pide al estudiante identifique la diferencia entre la variable horizontal del histograma y una función de densidad empírica, se observó que la gran mayoría de los estudiantes de Ingeniería Comercial no contestó (71.4%), mientras que un 21.4% afirmó que no son iguales las dos curvas. Respecto a los estudiantes de Medicina, un 50% consideró que ambas curvas presentan diferencias. Sin embargo, el problema, en ambos casos surgió a la hora de justificar, pues en ambos grupos la mayoría de los estudiantes dio argumentos que no confirmaban la afirmación dada. Esta es un área de oportunidad para la didáctica sobre cómo hacer transitar a los estudiantes de la variable estadística a la variable aleatoria. Ruiz, Batanero y Arteaga (2011) muestran que son pocos los estudiantes que son capaces de generalizar los resultados obtenidos con la variable estadística para realizar una inferencia informal sobre la variable aleatoria.

## ■ **Discusión**

Desde la perspectiva actual, la enseñanza de la Probabilidad y Estadística recomienda desarrollar el pensamiento y razonamiento estadístico (Mallows, 1998; Wild y Pfannkuch, 1999) para formar un estudiante crítico (Behar, 2018) y capaz de resolver problemas de modelación de la realidad y en la toma de decisiones (Ruiz y Albert, 2013). En este proceso, la variable toma un papel fundamental.

Desde la dimensión epistemológica la variable tiene gran complejidad no sólo porque en Probabilidad y Estadística toma diversos significados en el proceso de modelación (variable de interés, variable de observación) y por su naturaleza empírica, la variable estadística, y teórica, la variable aleatoria, sino también porque arrastra consigo, la complejidad venida del álgebra y el cálculo (Ursini y Trigueros, 2006).

Desde la dimensión didáctica, a pesar de que desde su formación temprana escolar, los estudiantes tratan con el concepto de variable, en rara ocasión reciben una instrucción formal acerca de sus diversos significados y diferencias. De esta forma, la variable pasa prácticamente inadvertida desde el Álgebra, luego por ecuaciones y funciones, al Análisis, donde incluso derivan e integran funciones. Al llegar a los cursos de Estadística y Probabilidad, los estudiantes se enfrentan a significados adicionales como lo son la variable estadística y la variable aleatoria que, además, es una función (Ruiz, Albert y Batanero, 2006). Los modelos de probabilidad están definidos sobre ciertos conceptos de variables muy específicos, vinculados al azar y aleatoriedad, que no los hacen coincidentes entre el tipo de variable matemática y en probabilidad y estadística.

Desde la dimensión cognitiva, diversos autores reportan dificultades específicas sobre la variable en su contexto algebraico (Escalante y Cuesta, 2012, entre otros). Este trabajo muestra que su complejidad también ha de ser vista de forma integral, desde los distintos papeles y significados que toma en el proceso de razonamiento y pensamiento estadístico, así como los contextos educativos específicos. En ese estudio se muestra una clara diferencia en las respuestas que son capaces de dar los estudiantes de Medicina y de Ingeniería Comercial. Los estudiantes de ambos grupos pueden entender un histograma puesto que en general, las respuestas asociadas a él son correctas. Sin embargo, no se reconoce la relación que puede haber con la curva (suave) que se asocia al histograma. Las comparaciones entre ellas que hacen los estudiantes son diversas e, incluso, dan a entender que no entienden qué es y qué representa cada gráfico. Llama la atención la diferencia de interpretación de ciertas preguntas por parte de ambos grupos de estudiantes, considerando que ambos han tenido acceso a la misma información e incluso con el mismo profesor. Varios estudiantes de Ingeniería Comercial responden a preguntas que son objetivas, con comentarios de percepción. Por ejemplo, "Razón ya que los datos dependen de la percepción de cada persona" es una respuesta no poco común cuando se pregunta por Escala de Medición. O bien: "Solo centrarse en pagar adicional y no tener la posibilidad de disminuir el precio anterior"; "que los datos tienen un signo (+)"; "el problema es que el fabricante no sabe cuánto puede modificar el precio de su producto" son algunas de las respuestas cuando se pide identificar el problema en el caso del fabricante de Dentífrico, del problema 3 de la segunda parte. Esto no se observa en los estudiantes de Medicina. Cabe señalar que la carrera de Ingeniería Comercial tiene un perfil más bien humanístico, aunque contiene una fuerte componente cuantitativa, mientras que medicina es considerada una carrera científica.

No se descarta que dichas diferencias entre ambos grupos estén relacionadas con las puntuaciones que se solicitan en la Prueba de Selección Universitaria para ingresar a ambas carreras. En el sistema de Educación Superior chileno, los estudiantes deben rendir una Prueba de Selección Universitaria, que es una prueba estandarizada, en la cual pueden obtener como máximo 850 puntos. Esta prueba estandarizada es la misma para todos los candidatos a ingresar a las universidades de ese país. El puntaje de admisión en las diferentes carreras lo define cada universidad, en base a estándares fijados por el Ministerio de Educación y la demanda local. Para el nuevo ingreso Ingeniería Comercial se requiere un puntaje mínimo de 475 puntos, y en la carrera de medicina, 600. De esta forma, se espera que los estudiantes de Medicina lleguen a la Universidad con una mejor preparación y se espere de ellos un mejor desempeño que los estudiantes de Ingeniería Comercial. Sin embargo, la discusión sobre las diferencias de las perspectivas de ambos grupos sigue siendo válida, puesto que la diferencia entre las respuestas de ambos grupos es muy patente.

## ■ Conclusiones

El concepto de variable tiene diversos significados por los que un estudiante ha de transitar a lo largo de su paso por el sistema escolar, desde el álgebra, el cálculo, hasta probabilidad y estadística y otras disciplinas científicas como física o química. Además, la variable forma parte del proceso del planteamiento del problema hasta su solución e interpretación, por lo que su complejidad no sólo es científica sino metodológica.

Este trabajo alerta a la comunidad de investigación educativa sobre lo que en un principio pareciera algo casi obvio, como lo es el concepto de variable y que, sin embargo, resulta de una gran complejidad epistémica, didáctica y cognitiva, más cuando es visto desde perspectivas tan diferentes como la algebraica y la estadística. La variable juega un papel decisivo en el proceso de modelación matemática y estocástica para resolver problemas del entorno y profesionales, así como desde una perspectiva integral ya que interviene desde el proceso de identificación del problema hasta el de modelación e interpretación de resultados. Es por ello que es importante hacer estudios que hagan visible este problema.

## ■ Referencias bibliográficas

- Albert, J. A. y Ruiz, B. (2014). Dificultades en estudiantes universitarios del estadístico como variable aleatoria en la distribución del muestreo de medias. *Acta Latinoamericana de Matemática Educativa* 27, 995-1003.
- Behar, R. (2018). Importancia del contexto en la formación del pensamiento y la cultura estadística. Bogotá, Colombia: *Tercer encuentro colombiano de educación estocástica*, 88-110.
- Cortinez, A., Alamilla, N., Albert, J. A. y Ríos, J. (2015). Razonamiento acerca del significado de los parámetros en los modelos de probabilidad en estudiantes universitarios. *Acta Latinoamericana de Matemática Educativa* 28, 283-290.
- Daniel, W. y Cross, C. (2018). *Biostatistics: a foundation for analysis in the health sciences*. New York: John Wiley & Sons.
- Escalante, J. y Cuesta, A. (2012). Dificultades para comprender el concepto de variable: un estudio con estudiantes universitarios. *Educación Matemática*, 24 (1): 107-132.
- Figueroa, S. y Aznar, M. (2017). Significados personales sobre la vinculación entre una variable estadística y su variable binomial asociada en el contexto de un problema. 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*.
- Herrera, H., Cuesta, A. y Escalante, J. (2016). El concepto de variable: un análisis con estudiantes de bachillerato. *Educación matemática*, 28(3), 217-240.
- Newbold, P., Carlson y W. Thorne, B.. (2008). *Estadística para administración y economía* (6ª ed.). Madrid: Pearson Educación.
- Mallows, C. (1998). The Zeroth Problem. *American Statistician*, 52, 1-9.
- Ruiz, B., Albert J. y Batanero, C. (2006). An exploratory study of students' difficulties with random variables. En A. Rossman y B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador (Bahía): International Association for Statistical Education.
- Ruiz, B., Batanero, C. y Arteaga, P. (2011). Vinculación de la Variable Aleatoria y Estadística en la Realización de Inferencias Informales por parte de Futuros Profesores. *Bolema*, 24, 431-449.
- Ruiz, B. (2013). *Análisis epistemológico de la variable aleatoria y comprensión de objetivos matemáticos relacionados por estudiantes universitarios*. Tesis doctoral. Universidad de Granada, España.
- Ruiz, B. y Albert, J. A. (2013). La relación entre la variable aleatoria y la variable estadística: un análisis epistemológico disciplinar. En J. M. Contreras, G. R. Cañadas, M. M. Gea y P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria* (pp. 383-390). Granada, Departamento de Didáctica de la Matemática de la Universidad de Granada, 2013.
- Ursini, S., Trigueros, M. y Lozano, D. (2000). La conceptualización de la variable en la enseñanza media. *Educación matemática*, 12(02), 27-48.
- Ursini, S. y Trigueros, M. (2006). ¿Mejora la comprensión del concepto de variable cuando los estudiantes cursan matemáticas avanzadas? *Educación Matemática*, 18, N° 3, 5-38.
- Wild, C. J., y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International statistical review*, 67(3), 223-248.