Universidad de Granada

Programa de Doctorado en Ciencias de la Educación

Análisis epistemológico y evaluación de la comprensión del concepto de variable aleatoria en estudiantes de secundaria chilenos

Valeria Constanza Bizet Leyton

Tesis doctoral

Departamento de Didáctica de la Matemática

Universidad de Granada

Granada, 2023

Editor: Universidad de Granada. Tesis Doctorales
Autor: Valeria Constanza Bizet Leyton
ISBN: 978-84-1195-238-5
URI: https://hdl.handle.net/10481/90703

Esta investigación fue financiada principalmente por el Gobierno de Chile mediante una Beca de Doctorado en el Extranjero (folio 72200367) y parcialmente por una beca de Movilidad Internacional de Estudiantes del Programa de Doctorado de la Universidad de Granada, para llevar a cabo una pasantía de tres meses en la Sam Houston State University de Estados Unidos.

Agradecimientos

A mis directores de tesis, por aceptarme como su doctoranda en la Universidad de Granada y embarcarse en el desarrollo de esta tesis. Además, por todo lo que me han enseñado durante este proceso y confiar en mi trabajo. Ambos han aportado en la construcción de mi perfil investigador que he logrado desarrollar hasta ahora.

Al doctor Dustin Jones por recibirme en la Sam Houston State University durante mi pasantía, gracias por compartir conmigo su trabajo y ayudarme en algunas etapas de la investigación. A los doctores Audy Salcedo y Claudia Vásquez por su buena disposición en la revisión de esta tesis y su valiosa valoración. A las personas que participaron en esta investigación, puntualmente a los docentes universitarios que cedieron una sesión de su asignatura para aplicar el cuestionario y los estudiantes que respondieron este instrumento.

A mi familia. En especial a mi madre Nancy por inculcarme valores como el respeto, la honestidad y la perseverancia, gracias por enseñarme a luchar para cumplir mis sueños respetando a los que me rodean y apoyarme incondicionalmente A mi tía Inés y hermana Valentina, por acompañarme a la distancia y ayudarme siempre que necesito. A mi sobrino Jorgito, por inspirarme cada día a ser mejor persona y profesional. A David por su amor, compañía y brindar su ayuda cada vez que necesito.

A mis amigos. En especial a Norma por su apoyo y su compañía en los buenos y malos momentos que he vivido en España. También a Casandra, por haber hecho más amena mi estancia en Estados Unidos, recibirme en su casa y considerarme un integrante más de su familia.

A Dios por guiar mi vida y poner en mi camino personas con mis mismos valores.

Resumen

En esta investigación se ha explorado la comprensión de variable aleatoria y distribuciones binomial y normal en el contexto escolar chileno, centrando la atención en el análisis del tratamiento de los temas de interés en el currículo escolar y libros de texto chilenos. Así como también en la evaluación de la comprensión de dichos temas en estudiantes egresados de educación escolar chilena, debido a que en las escuelas de Chile los temas de interés son trabajados en los últimos grados de educación secundaria (entre los grados 10 y 12). Principalmente se utilizó como fundamento teórico algunas herramientas del Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos. En este sentido, la presente tesis doctoral se organizó en seis estudios, donde se abordaron la indagación de antecedentes, un análisis de libros de texto y de la normativa curricular chilena, el diseño y validación de un instrumento de evaluación, la valoración de la comprensión en estudiantes chilenos.

En el Estudio 1, se ha desarrollado una revisión bibliográfica sobre la enseñanza y aprendizaje de variable aleatoria y distribuciones binomial y normal. Específicamente fueron indagadas las investigaciones más relevantes en torno a: i) la enseñanza de los temas de interés, donde se abordaron sus aspectos epistemológicos y didácticos, en el contexto escolar y universitario; y ii) el aprendizaje de los temas en cuestión, donde se trataron sus aspectos cognitivos en el ámbito escolar y universitario. Entre los resultados obtenidos se evidencia, que los estudios realizados desde un ámbito epistemológico han identificado de manera superficial elementos de la variable aleatoria (problemas, procedimientos, representaciones, etc.) necesarios para su enseñanza en la escuela y universidad, además no se han hallado estudios que entreguen directrices sobre los elementos principales de la distribución binomial a considerar en su enseñanza. En los estudios desarrollados desde una perspectiva cognitiva, se hallaron alrededor de la misma cantidad que abordaban la comprensión de variable aleatoria y su distribución de probabilidad en la educación escolar o universitaria. Sin embargo, en la literatura indagada existen vacíos que serían importantes investigar, como estudiar articuladamente el nivel de comprensión de variable aleatoria y los modelos de probabilidad al término de la educación escolar.

En el segundo y tercer estudio fue analizado el tratamiento de variable aleatoria y distribuciones binomial y normal en la normativa curricular chilena vigente durante el desarrollo de esta investigación y en libros de texto chilenos dirigidos a los grados 10 a 12 (15 a 18 años). Concretamente en el Estudio 2 se han identificado las situaciones-problemas sobre los temas de interés que promueven tanto el currículo escolar como los libros de texto indagados. Los resultados evidencian situaciones-problemas que son promovidas en el currículo escolar pero que en los libros indagados existe ausencia y mínima presencia de estas, como calcular probabilidades asociadas a distribuciones binomial y normal empleando una herramienta tecnológica, y diferenciar entre variables aleatorias y variables con dependencia funcional, respectivamente. También fue observado que hay situaciones-problemas fomentadas en los libros de texto analizados que no son sugeridas en el currículo chileno, tales como: diferenciar entre variables aleatorias discretas y continuas, calcular el valor de incógnitas de manera que la función propuesta sea de probabilidad, calcular probabilidades en una distribución normal estándar, representar en un gráfico la distribución normal y calcular los valores correspondientes a una probabilidad dada en el contexto de la distribución normal.

Asimismo, en el Estudio 3 se ha reconocido el lenguaje, los conceptos, las proposiciones, los procedimientos y los argumentos vinculados a los temas en cuestión, que sugieren los lineamientos curriculares y libros escolares chilenos. Los resultados muestran que: i) sobre el lenguaje relacionado con la variable aleatoria y distribución normal se evidenció que entre los lineamientos curriculares y los libros indagados existe congruencia, mientras que en el contexto de la binomial se observó desarmonía entre los documentos indagados; ii) en los conceptos vinculados a la variable aleatoria se ha evidenciado coherencia entre el currículo escolar y los libros de texto, a diferencia del contexto de los modelos de probabilidad donde se observaron discrepancias entre los documentos analizados; iii) respecto a las proposiciones ligadas a la distribución binomial se evidenció armonía entre los lineamientos curriculares y los textos escolares, en tanto que en el contexto de la variable aleatoria y distribución normal se identificaron ciertas incongruencias entre los documentos indagados; iv) sobre los procedimientos relacionados con la variable aleatoria y los modelos de probabilidad existió desarmonía entre el currículo escolar y libros analizados;

y v) en los argumentos vinculados a la variable aleatoria, distribución binomial y distribución normal se encontraron incongruencias entre los documentos indagados.

En el cuarto y quinto estudio se ha construido y validado un instrumento orientado a evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en el contexto escolar chileno. En el Estudio 4, primero se diseñó una guía de situaciones-problemas sobre los temas de interés a partir del análisis de lineamientos curriculares chilenos (GSP-VADP). Luego fue seleccionado un conjunto inicial de ítems representantes de las situaciones-problemas que componen la guía, estos fueron elegidos de investigaciones previas o libros de texto anteriormente analizados. Posteriormente se realizó la validez de contenido de aquellos ítems por medio de la valoración de un grupo de expertos, cuya evaluación permitió seleccionar un conjunto final de ítems. Los resultados muestran que la GSP-VADP está constituida por diversas situaciones-problemas desde las que emergen los restantes objetos matemáticos primarios. Esta herramienta es viable utilizarla para identificar ítems representantes de sus situaciones-problemas por medio de juicio de expertos, debido a que el conjunto final de ítems obtuvo un coeficiente de validez y concordancia bueno (0,87).

En tanto que en el Estudio 5 fue analizada la validez de constructo y fiabilidad del instrumento, donde la versión inicial del cuestionario se aplicó en una muestra dirigida de 80 estudiantes egresados de educación escolar chilena. La validez de constructo del instrumento se ha indagado mediante un análisis factorial exploratorio (AFE) y confirmatorio (AFC), y su fiabilidad empleando el coeficiente de alfa de Cronbach. Los resultados muestran que, en el AFE, la estructura empírica del cuestionario está compuesta por seis factores que explican el 58% de la varianza total, cada uno relacionado con un campo de problema (constituido por tres a cuatro ítems) sobre los temas en cuestión: variable aleatoria, función de probabilidad y función de distribución de una variable aleatoria discreta, valores de posición central o de dispersión asociados a una variable aleatoria, distribuciones binomial y normal. En el AFC, la mayoría de los coeficientes factor-ítem poseen valores adecuados y la estructura del instrumento tiene medidas de ajuste aceptables, además el alfa de Cronbach de 0,824 demuestra que el cuestionario es fiable y existe consistencia entre los ítems y el constructo. Por tanto, la versión final del cuestionario (19 ítems) posee evidencias que es un instrumento válido y fiable para indagar la compresión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar chilena.

En el Estudio 6 fue evaluada la comprensión de variable aleatoria y distribuciones binomial y normal en 101 estudiantes egresados de educación escolar chilena, por medio de la aplicación del cuestionario anteriormente validado. Los resultados evidencian sobre la comprensión de variable aleatoria, que la mayoría de las situaciones-problemas valoradas en el cuestionario han sido resueltas por los participantes, pero en un bajo porcentaje. Aún más limitados estudiantes de la muestra han logrado emplear los distintos objetos matemáticos involucrados en la solución de las tareas evaluadas. En consecuencia, los estudiantes egresados de educación escolar chilena poseen una comprensión baja de variable aleatoria. En cuanto a la comprensión de los modelos de probabilidad, todas las situaciones-problemas evaluadas en el cuestionario han sido resueltas por los participantes, aunque en una baja proporción. También, pocos participantes lograron utilizar la diversidad de objetos matemáticos primarios asociados a la resolución de las situaciones-problemas sobre distribución binomial. Asimismo, pocos estudiantes de la muestra lograron aplicar algunos objetos primarios asociados a la solución de las tareas en torno a distribución normal. Por consiguiente, los participantes poseen una comprensión baja de distribución binomial y distribución normal.

Finalmente, entre las principales aportaciones de la presente tesis doctoral se destacan: i) la información relacionada con las tareas sobre variable aleatoria, modelo binomial y modelo normal y los diversos objetos matemáticos primarios involucrados en su solución, que son promovidos en los textos escolares chilenos; ii) la herramienta, Guía de Situaciones-Problemas sobre Variable Aleatoria y sus Aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar Chileno; iii) el instrumento, cuestionario para evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en estudiantes egresados de educación escolar; y v) los resultados de la evaluación articulada sobre la comprensión de variable aleatoria y modelos de probabilidad en estudiantes egresados de educación escolar chilena.

Summary

This study delves into the comprehension of random variables, binomial and normal distributions within the context of Chilean schools. The primary focus is analyzing how these subjects are addressed in the Chilean school curriculum and textbooks. Additionally, it aims to evaluate the understanding of these topics among students who have completed their education in Chilean schools, given that these subjects are typically covered in the latter grades of secondary education (between grades 10 and 12). The theoretical framework for this research relies mainly on the Ontosemiotic Approach to Mathematical Knowledge and Instruction. The doctoral thesis is structured into six studies, encompassing background research, examining textbooks and Chilean curricular regulations, developing and validating an assessment tool, and evaluating students' comprehension in the Chilean educational context.

Study 1 involved conducting a comprehensive literature review on the instruction and learning of random variables and binomial and normal distributions. This review specifically focused on the most pertinent research related to i) the teaching of these subjects, addressing both epistemological and didactic aspects in both school and university settings, and ii) the learning of these topics, examining cognitive aspects within the school and university context. The findings from the literature review indicate that studies conducted from an epistemological standpoint have only superficially identified elements of the random variable (such as problems, procedures or representations) crucial for effective teaching at both school and university levels. Additionally, studies that offer guidance on the key elements of teaching the binomial distribution have yet to be identified. In studies conducted from a cognitive perspective, similar investigations were found addressing the understanding of random variables and their probability distributions in school or university education. However, notable gaps exist in the literature, warranting further exploration. One such gap pertains to the need to investigate the level of understanding of random variables and probability models among students upon completion of their school education.

The second and third studies examine how random variables, as well as binomial and normal distributions, are addressed in the current Chilean curricular regulations and textbooks intended for students in grades 10 to 12 (ages 15 to 18) during the timeframe of

this research. Specifically, Study 2 identified the problem situations related to the subjects of interest promoted by the school curriculum and the textbooks under investigation.

The findings reveal that certain problem situations are endorsed in the school curriculum but are either absent or minimally covered in the surveyed textbooks. Examples include the calculation of probabilities associated with binomial and normal distributions using technological tools and distinguishing between random variables and variables with functional dependence. Additionally, observations indicated instances where textbooks introduce situations or problems not suggested in the Chilean curriculum. These include differentiation between discrete and continuous random variables, calculating the value of unknowns to establish a probability function, determining probabilities in a standard normal distribution, graphically representing the normal distribution, and calculating values corresponding to a given probability within the context of the normal distribution.

Similarly, in Study 3, the examination focused on recognizing the language, concepts, propositions, procedures, and arguments associated with the topics under consideration, as suggested by Chilean curricular guidelines and school textbooks. The results reveal the following: (i) Concerning the language related to the random variable and normal distribution, it was evident that there is congruence between the curricular guidelines and the surveyed books. However, discrepancies were observed between the analyzed documents in the context of the binomial distribution. (ii) In terms of concepts linked to the random variable, coherence was evident between the school curriculum and the textbooks. Conversely, inconsistencies were observed between the documents analyzed in the context of probability models. (iii) Regarding propositions linked to the binomial distribution, there was harmony between the curricular guidelines and the textbooks. However, certain incongruities were identified between the surveyed documents in the context of the random variable and normal distribution. (iv) Concerning procedures related to the random variable and probability models, there was disharmony between the school curriculum and the analyzed books. (v) In the arguments associated with the random variable, binomial distribution, and normal distribution, incongruities were found between the surveyed documents.

In the fourth and fifth studies, an assessment instrument was developed and validated to evaluate the comprehension of random variables and binomial and normal distributions within the Chilean school context. In Study 4, a guide for problem situations about interest was initially crafted based on the analysis of Chilean curricular guidelines (GPS-RVPD). Subsequently, an initial set of items representing the problem situations outlined in the guide was chosen from prior research or previously examined textbooks. Following this, the content validity of these items was evaluated by a panel of experts, whose assessment led to the final selection of items. The results indicate that the GPS-RVPD encompasses various problems, giving rise to the remaining primary mathematical components. The tool is deemed suitable for identifying items that represent its problem situations through expert judgment, as the final set of items achieved a favorable validity and agreement coefficient of 0,87.

Study 5 examined the instrument's construct validity and reliability, involving administering the initial version of the questionnaire to a targeted sample of 80 Chilean school graduates. The instrument's construct validity was assessed through both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), while reliability was gauged using Cronbach's alpha coefficient. The EFA findings revealed that the questionnaire's empirical structure consists of six factors, explaining 58% of the total variance. Each factor is associated with a specific problem domain, comprising three to four items related to the topics under consideration: random variable, probability function and distribution function of a discrete random variable, values of central position or dispersion associated with a random variable, binomial, and normal distributions. In the CFA, most of the factor-item coefficients demonstrated adequate values, and the overall structure of the instrument exhibited acceptable measures of fit. Furthermore, a Cronbach's alpha coefficient of 0.824 indicated that the questionnaire is reliable, demonstrating consistency between the items and the construct. Consequently, the final version of the questionnaire (consisting of 19 items) provides evidence of being a valid and reliable tool for investigating the understanding of random variables and probability distributions among Chilean school graduates.

In Study 6, the assessment of the understanding of random variables and binomial and normal distributions was conducted among 101 Chilean school graduates using the previously validated questionnaire. The findings indicate that the participants addressed most of the problem situations evaluated in the questionnaire, albeit with a low success rate. Furthermore, a limited percentage of students in the sample demonstrated the ability to utilize the various mathematical components essential for solving the assessed tasks. Consequently, Chilean school leavers must demonstrate a greater understanding of random variables. In terms of comprehending probability models, the participants resolved all the problem situations assessed in the questionnaire, though at a relatively low proportion. Additionally, only a few participants successfully employed the diverse primary mathematical components of solving problem situations related to the binomial distribution. Similarly, a few students in the sample applied some primary objects linked to the resolution of tasks involving the normal distribution. Hence, participants displayed a limited understanding of both binomial and normal distribution.

Finally, the primary contributions of this dissertation encompass (i) insights into the tasks related to random variables, the binomial model, and the normal model, as well as the various primary mathematical components involved in their solution, as promoted in Chilean school textbooks; (ii) the development of the tool, "Guía de Situaciones-Problemas sobre Variable Aleatoria y sus Aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar Chileno" (Guide of Situations-Problems on Random Variable and its Applications in Probability Distributions according to the Chilean School Curriculum); (iii) the creation of the assessment instrument, a questionnaire to evaluate the understanding of random variables, and binomial and normal distributions in school leavers; and (v) the presentation of results from the comprehensive evaluation of the understanding of random variables and probability models among Chilean school leavers.

Índice

INTRODUCCIÓN GENERAL

Esta investigación principalmente se centra en la comprensión de variable aleatoria y distribuciones binomial y normal en estudiantes egresados de educación escolar chilena. El contexto es elegido debido a que en las escuelas de Chile los temas de interés son trabajados en los últimos grados de educación secundaria (entre los grados 10 y 12). En este sentido, la presente tesis doctoral se organiza en seis capítulos, donde se expone el planteamiento del problema, los antecedentes, un análisis de libros de texto y normativa curricular chilena, el diseño y validación de un instrumento de evaluación, la valoración de la comprensión en estudiantes chilenos y las conclusiones generales, además de un apartado de referencias y otro de anexos.

En el Capítulo 1, planteamiento del problema, inicialmente es justificada la importancia de la variable aleatoria y distribuciones de probabilidad. Además, se describe la manera que la normativa curricular chilena y lineamientos internacionales (Estados Unidos) trabajan los temas de interés. Posteriormente son presentadas las principales herramientas del marco teórico que sustenta esta tesis, y seguidamente se establecen los objetivos principales de la investigación y las hipótesis iniciales fundamentadas en estudios previos. También se expone la organización de esta tesis doctoral y un resumen de los principales aspectos metodológicos utilizados para su desarrollo.

En el Capítulo 2, antecedentes, se desarrolla una revisión bibliográfica que fundamenta los demás estudios incluidos en la presente tesis doctoral. Esto queda recogido en el Estudio 1. What does the research tell us about the understanding of the random variables and its probability distributions? (revista BEIO). Puntualmente son indagadas las investigaciones más relevantes en torno a la enseñanza, en el ámbito escolar y universitario, de la variable aleatoria y distribuciones binomial y normal, donde se abordan sus aspectos epistemológicos y didácticos. Asimismo, son analizadas las investigaciones más importantes sobre el aprendizaje, en el contexto escolar y universitario, de los temas de interés, allí se tratan sus aspectos cognitivos. También se incluye un apartado de investigaciones excluidas en la publicación de BEIO, debido a que en la búsqueda de la literatura se integraron nuevos descriptores y extendió el límite de año de publicación.

En el Capítulo 3, se lleva a cabo el análisis de la normativa curricular chilena vigente durante el desarrollo de esta tesis doctoral y la indagación de libros de texto chilenos dirigidos a los grados 10 a 12 (15 a 18 años), centrándose en la identificación de los objetos matemáticos primarios ligados a la variable aleatoria y distribuciones binomial y normal. Aquel trabajo queda recogido en dos estudios. En el Estudio 2. Situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad según libros de texto chilenos (revista Profesorado), se presenta las situaciones-problemas sobre los temas de interés que son reconocidas en el currículo escolar y libros de texto chilenos. Mientras que en el Estudio 3. Objetos matemáticos ligados a la variable aleatoria y sus aplicaciones sobre distribuciones de probabilidad en libros de texto de educación secundaria chilena (revista PNA), se expone el lenguaje, los conceptos, las proposiciones, los procedimientos y los argumentos vinculados a los temas en cuestión, que son reconocidos en los lineamientos curriculares y libros escolares chilenos.

En el Capítulo 4, se realiza la construcción de un instrumento y aproximación a su validez de contenido y de constructo, lo cual queda recogido en dos estudios. En el Estudio 4. Elaboración de una guía de situaciones-problema sobre variable aleatoria y sus aplicaciones a partir del currículo escolar chileno (revista Educación Matemática), inicialmente es abordado el diseño de una guía de tareas en torno a los temas de interés a partir del análisis de lineamientos curriculares chilenos y extranjeros (Estados Unidos). Posteriormente, se presenta la selección de un conjunto inicial de ítems representantes de las situaciones-problemas que componen la guía, desde la indagación de investigaciones previas y análisis de libros de texto. Finalmente es realizada la validez de contenido de aquellos ítems por medio de la valoración de un grupo de expertos, donde su evaluación permite seleccionar un conjunto final de ítems.

En tanto que en el Estudio 5. Cuestionario para valorar la comprensión de variable aleatoria y distribuciones de probabilidad en educación escolar (en evaluación), es desarrollada la validez de constructo del instrumento con una muestra de 80 estudiantes egresados de educación escolar chilena. Específicamente, mediante un análisis factorial exploratorio es indagada la estructura de la versión inicial del instrumento y por medio de un análisis factorial confirmatorio se realiza una aproximación a confirmar la estructura de la versión final del instrumento. También es analizada la fiabilidad de la versión final del instrumento por medio del método de consistencia interna.

En el Capítulo 5, se presenta el Estudio 6. Comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar chilena, donde aquella comprensión se evalúa en 101 estudiantes universitarios mediante la aplicación de un cuestionario previamente validado. Específicamente se dan a conocer los resultados obtenidos del análisis de las respuestas a cada uno de los ítems, donde se describen las categorías de respuestas e indagan los porcentajes de cada una. Asimismo, es reconociendo tanto el lenguaje, conceptos, procedimientos, proposiciones y argumentos que los estudiantes lograron emplear para solucionar las situaciones-problemas propuestas o avanzar parcialmente en su solución, como aquellos objetos matemáticos primarios que evidenciaron no manipular en la resolución de las tareas. Luego, se expone una síntesis de los resultados, incluyendo algunos estadísticos descriptivos de la puntuación conseguida por los participantes en cada ítem y en el cuestionario**.**

En el Capítulo 6 se exponen las conclusiones finales de esta tesis doctoral, donde primero es presentada una discusión sobre el logro de los objetivos e hipótesis de investigación planteadas en el Capítulo 1. Posteriormente se muestra una síntesis de las principales contribuciones de la tesis para la Educación Estadística. Finalmente son analizadas las limitaciones de esta investigación y presentadas algunas líneas para futuros trabajos.

CAPÍTULO 1: PLANTEAMIENTO DEL PROBLEMA

1.1 Introducción

En este capítulo de la presente tesis doctoral es planteado el problema de investigación en torno a la comprensión de variable aleatoria y distribuciones binomial y normal en estudiantes de educación escolar. Primero se justifica la importancia de la variable aleatoria y distribuciones de probabilidad, tanto en la formación del estudiante para la vida como en el estudio de contenidos de estadística posteriores. Después se describe la forma que la normativa curricular chilena y lineamientos internacionales abordan los temas señalados.

A continuación, son expuestas las principales herramientas del marco teórico que sustenta esta tesis. Posteriormente se establecen los objetivos principales de la investigación. Además, son propuestas las hipótesis iniciales fundamentadas en estudios previos. Finalmente se presenta la organización de esta tesis doctoral y un resumen de los principales aspectos metodológicos empleados para su desarrollo.

1.2 Importancia de la variable aleatoria y distribuciones de probabilidad

La enseñanza-aprendizaje de la probabilidad es fundamental tanto para comprender otros tópicos como para colaborar en la formación de los estudiantes para la vida, pues el azar y los fenómenos aleatorios envuelven nuestro entorno. Según Gal (2005) profesionales y adultos en cualquier ámbito de la vida deben interpretar o actuar en situaciones que incluyen elementos probabilísticos o estimar probabilidades de sucesos, por ello es necesario fomentar el desarrollo de la alfabetización probabilística (probability literacy). Esta es definida como "*la capacidad de acceder, utilizar, interpretar y comunicar información e ideas relacionadas con la probabilidad, con el fin de participar y gestionar eficazmente las demandas de las funciones y tareas que implican incertidumbre y riesgo del mundo real*" (Gal, 2012, p.4).

La alfabetización probabilística posee componentes que interactúan entre sí. Puntualmente las componentes cognitivas incluyen (Gal, 2005): i) las grandes ideas como la aleatoriedad, variación y previsibilidad/ incertidumbre; ii) el cálculo de probabilidades; iii) el lenguaje; iv) el contexto; y v) las preguntas críticas. Mientras que las componentes en torno a la disposición consideran: (Gal, 2005): i) la postura crítica; ii) las creencias y actitudes; y iii) los sentimientos personales sobre la incertidumbre y riesgo. En esta perspectiva Jones et al. (2007) incorporan a las grandes ideas en probabilidad las distribuciones de probabilidad, abarcando desde el concepto de variable aleatoria hasta algunos modelos de probabilidad (binomial, uniforme, normal y chi-cuadrado).

En la educación estocástica, la variable aleatoria es un tópico esencial dada su importancia en la teoría de la probabilidad y en la práctica estadística (Heitele, 1975). Su estudio involucra diversidad de elementos que se relacionan entre sí, tales como (Soong, 2004): aleatoriedad, experimento aleatorio, espacio muestral, álgebra de sucesos, suceso aleatorio, asignación de probabilidades (enfoque axiomático, clásico y frecuencial), función, distribución de probabilidad (también denominada función de probabilidad), media (o esperanza), desviación estándar y varianza, por lo que su comprensión puede entramar cierta complejidad.

Heitele (1975) incluye la variable aleatoria entre las ideas estocásticas fundamentales debido a que posee un rol importante en la evolución del cálculo de probabilidades, se presenta de manera implícita en diversas situaciones aleatoria y no es intuitiva en ciertas de sus cualidades. Además, señala que su distribución de probabilidad posee dos importantes características como la media y la desviación estándar, ya que en la sociedad actual permiten a la persona enfrentarse a datos estadísticos de forma crítica. Para Batanero (2004) la variable aleatoria es un concepto importante, pues introduce el análisis matemático en la probabilidad y se manifiesta en experiencias cotidianas como en juegos de azar, tiempo de espera en una fila y peso de una persona.

También afirma que esta posee tres ideas básicas: su distribución de probabilidad, su media y varianza.

Burrill y Biehler (2011) sugieren las distribuciones de probabilidad de variables aleatorias como un tema importante en estadística, pues permiten vincular esta disciplina con experiencias del mundo real, enseñar la estructura de la disciplina y lograr una comprensión profunda del tema a medida que los estudiantes maduran su conocimiento sobre estadística. En tanto que Batanero et al. (2016) proponen entre las ideas probabilísticas fundamentales las distribuciones de probabilidad, incluyendo algunos modelos de probabilidad como la distribución binomial y distribución normal, debido a su utilidad en la inferencia estadística. Según Rossman (2008) los modelos de probabilidad son la base para realizar inferencia estadística y particularmente Lee (2016) considera a aquellos entre los conceptos probabilísticos fundamentales para enseñar un enfoque de muestreo repetido.

Todos aquellos argumentos expuestos anteriormente justifican el interés e importancia de investigar la comprensión sobre variable aleatoria y sus modelos de probabilidad por parte de estudiantes escolares chilenos.

1.3 Normativa curricular

En este apartado es analizado el tratamiento de la variable aleatoria y sus modelos de probabilidad en normativas curriculares, para completar la información sobre su importancia en el ámbito escolar. Primero se exponen la indagación de la normativa curricular chilena, debido a que el estudio de valoración de esta tesis se realiza en el contexto chileno. Posteriormente es presentado el análisis de los estándares americanos y los lineamientos para la evaluación y enseñanza en educación estadística, pues son documento de referencia internacional para afrontar los desafíos de la educación estadística a nivel escolar.

1.3.1 Normativa curricular chilena

En Chile, la Ley General de Educación (20.370) guía la educación escolar obligatoria, la cual tiene una duración de trece años organizada en tres niveles educativos (ver Tabla 1): educación parvularia compuesta por un curso escolar (kínder); educación básica integrada por ocho cursos consecutivos (1° a 8° básico); y educación media constituida por cuatro cursos contiguos (1° a 4° medio). Sin embargo, a partir del año 2027 se integrará al último nivel señalado los cursos de 7° y 8° básico. También la Ley General de Educación norma los principales documentos del currículo nacional, denominados Bases Curriculares y Programas de Estudio, que son elaborados por el Ministerio de Educación (MINEDUC). En el primer documento nombrado, por asignatura se establecen los aprendizajes mínimos a lograr en cada curso escolar, mientras que en el segundo documento se lleva a cabo una organización temporal de un año académico, sobre los aprendizajes correspondientes a cada asignatura por curso.

Tabla 1

Correspondencia entre cursos escolares chilenos y edad

Edad	Curso escolar	Nivel de educación
5-6 años	kínder	Parvularia
6-7 años	1° básico	Básica
7-8 años	2° básico	
8-9 años	3° básico	
9-10 años	4° básico	
10-11 años	5° básico	
11-12 años	6° básico	
12-13 años	7° básico	
13-14 años	8° básico	
14-15 años	1° medio	Media
15-16 años	2° medio	
16-17 años	3° medio	
17-18 años	4° medio	

Nota. Elaboración propia.

Las Bases Curriculares en la asignatura de matemática organizan los conocimientos en torno a ejes temáticos: desde 1° a 6° básico, cada curso incluye cinco ejes, números y operaciones, patrones y álgebra, geometría, medición, y datos y probabilidades; a partir de 7° básico hasta 4° medio, cada curso abarca cuatro ejes, números, álgebra y funciones, geometría, y probabilidad y estadística.

1.3.1.1 Elementos relacionadas a la variable aleatoria en las Bases Curriculares

La educación escolar chilena desde los primeros cursos introduce elementos vinculados de manera implícita a la variable aleatoria y su distribución de probabilidad (ver Tabla 2). En el eje datos y probabilidades, la idea de aleatoriedad se aborda a partir de 2° básico mediante el lanzamiento de objetos físicos como dados o monedas, observación y registro de su resultado. Luego en 4° básico se incorpora el uso de algún software para simular aquellos lanzamientos. El concepto de experimento aleatorio es introducido en 4° básico y formalizado en 6° básico, en estos cursos escolares es utilizado tanto material concreto (naipes, monedas y dados cúbicos o de otra forma regular) como softwares para simularlo, y se requiere realizar la distinción entre este tipo de experimentos y experimentos deterministas.

Otros elementos de interés son los gráficos de barras y las tablas, debido a que ambos permiten representa la distribución de probabilidad de una variable aleatoria, además de las tablas de conteo que posibilitan representar variables aleatorias (discretas). A partir de los primeros cursos de educación básica se sugiere su estudio, por ejemplo, en 2° básico y 3° básico es necesario registran resultados de experimentos aleatorios en tablas de conteo y gráficos de barras. También sobre estos últimos se requiere su construcción, lectura e interpretación, con la finalidad de responder preguntas que conduzcan a realizar predicciones. Después en 4° básico es introducido el uso de software para la elaboración de aquellos elementos.

Tabla 2

Objetivos de aprendizaje relacionados con la variable aleatoria en las Bases Curriculares

Curso escolar	Objetivos de aprendizaje
2° básico	20. Recolectar y registrar datos para responder preguntas estadísticas sobre juegos con monedas y dados, usando tablas de conteo. 21. Registrar en tablas y gráficos de barras simple, resultados de juegos aleatorios con dados y monedas. 22. Construir, leer e interpretar gráficos de barras simple. (MINEDUC, 2018, p.232)
3° básico	24. Registrar y ordenar datos obtenidos de juegos aleatorios con dados y monedas, encontrando el menor, el mayor y estimando el punto medio entre ambos. 25. Construir, leer e interpretar gráficos de barras simple con escala, de acuerdo con información recolectada o dada. (MINEDUC, 2018, p.237)
4° básico	26. Realizar experimentos aleatorios lúdicos y cotidianos, y tabular y representar mediante gráficos de manera manual y/o con software educativo. 27. Leer e interpretar gráficos de barras simple con escala y comunicar sus conclusiones. (MINEDUC, 2018, p.244)
5° básico	23. Calcular el promedio de datos e interpretarlo en su contexto. 24. Describir la posibilidad de ocurrencia de un evento por sobre la base de un experimento aleatorio, empleando los términos: seguro, posible, poco posible, imposible. 25. Comparar probabilidades de distintos eventos sin calcularlas. (MINEDUC, 2018, p.249)
6° básico	23. Conjeturar acerca de la tendencia de resultados obtenidos en repeticiones de un mismo experimento con dados, monedas u otros, de manera manual y/o usando software educativo (MINEDUC, 2018, p.254)
7º básico	18. Explicar las probabilidades de eventos obtenidos por medio de experimentos de manera manual y/o con software educativo: -Estimándolas de manera intuitiva. -Utilizando frecuencias relativas. -Relacionándolas con razones, fracciones o porcentaje. 19. Comparar las frecuencias relativas de un evento obtenidas al repetir un experimento de forma manual y/o con software educativo, con la probabilidad obtenida de manera teórica, usando diagramas de árbol, tablas o gráficos.

	(MINEDUC, 2015, p.109)
8° básico	7. Mostrar que comprenden la noción de función por medio de un cambio lineal: -Utilizando tablas: -Representando de manera gráfica (plano cartesiano, diagramas de Venn), de manera manual y/o con software educativo. (MINEDUC, 2015, p.114) 17. Explicar el principio combinatorio multiplicativo: -A partir de situaciones concretas. -Representándolo con tablas y árboles regulares, de manera manual y/o con software educativo. -Utilizándolo para calcular la probabilidad de un evento compuesto (MINEDUC, 2015, p.115)
1° medio	14. Desarrollar las reglas de las probabilidades, la regla aditiva, la regla multiplicativa y la combinación de ambas, de manera concreta, pictórica y simbólica, de manera manual y/o con software educativo, en el contexto de la resolución de problemas. 15. Mostrar que comprenden el concepto de azar: -Experimentando con la tabla de Galton de manera manual y/o con software educativo. -Realizando análisis estadísticos, empezando por frecuencias relativas. -Utilizando probabilidades para describir el comportamiento azaroso. -Resolviendo problemas de la vida diaria y de otras asignaturas (MINEDUC, 2015, p.120)
2° medio	11. Utilizar permutaciones y la combinatoria sencilla para calcular probabilidades de eventos y resolver problemas (MINEDUC, 2015, p.125)

Nota. Elaboración propia.

Sobre el concepto de suceso aleatorio y la asignación de probabilidades, su estudio comienza en 5° básico donde se debe identificar el grado de posibilidad de ocurrencia de sucesos (seguro, posible, poco posible o imposible). En el siguiente curso escolar (6° básico) es abordado el enfoque frecuencial de probabilidad, dado que se requiere estimar la posibilidad de ocurrencia de sucesos aleatorios, a través de la frecuencia relativa observada en un gran número de repeticiones del experimento, utilizando material manipulativo o software. Luego en 7° básico se estudia el enfoque clásico de probabilidad, allí es necesario calcular la probabilidad de sucesos aleatorios, mediante la proporción entre el número de casos favorable de un suceso y número de casos posibles (donde todo suceso del espacio muestral tiene que ser igualmente probable), y se debe comparar la probabilidad de un suceso con su aproximación.

También se relaciona a los temas en cuestión: el concepto de promedio introducido en 5° básico, debido a que la esperanza de una variable aleatoria corresponde a la media de su distribución de probabilidad; la noción de función, pues la variable aleatoria y su distribución de probabilidad son funciones, aquella es abordada en 8 básico (eje álgebra y funciones) junto con sus representaciones en diagrama de Venn, tabla de valores, expresión algebraica y plano cartesiano.

La combinatoria es otra idea vinculada a los temas de interés y se estudia en el eje probabilidad y estadística, aunque en educación básica sólo se aborda el principio multiplicativo (regla del producto). Este es introducido en 8° básico, donde se requiere comprender, explicar y utilizarlo en contextos de la vida cotidiana para calcular la cardinalidad de un espacio muestral, además se debe utilizar el diagrama de árbol para representar problemas combinatorios y calcular probabilidades. Posteriormente, en educación media se profundiza el estudio de la combinatoria, específicamente en 1° medio es aplicado el principio multiplicativo en la resolución de problemas para determinar la cantidad de casos posibles en diversas situaciones que implican ordenamientos de objetos, con la finalidad de realizar cálculos sin necesidad de contar uno por uno los casos. En 2° medio se introducen las operaciones combinatorias de permutación, variación y combinación, con el propósito de emplearlas para el cálculo de probabilidades.

1.3.1.2 La variable aleatoria y sus modelos de probabilidad en la normativa curricular

De acuerdo con la normativa curricular chilena (ver Tablas 3 y 4), el eje probabilidad y estadística propone para 2° medio introducir la idea de variable aleatoria, centrándose en las de tipo discretas (sólo finitas). Puntualmente se requiere incluir: i) su diferenciación con la variable algebraica; ii) su definición en experimentos aleatorios y representación en tabla; iii) identificación de su dominio y recorrido; iv) asignación de probabilidades a algunos de sus valores (desde el enfoque clásico de probabilidad); v) determinación de su distribución de probabilidad (función de probabilidad) y representación de esta en tabla y gráfico de barras. En 3° medio, bajo el contexto de análisis de datos, son tratadas algunas medidas de dispersión como el rango, varianza, desviación media y desviación estándar, momentos en que se sugiere abordar la media, desviación estándar y varianza de una variable aleatoria discreta y realizar una relación entre la estadística y probabilidad. Para 4° medio, se trabajan las variables aleatorias continuas.

Tabla 3

Objetivos de aprendizaje sobre variable aleatoria y modelos de probabilidad en las Bases Curriculares

Curso escolar	Objetivos de aprendizaje
2° medio	10. Mostrar que comprenden las variables aleatorias finitas: -Definiendo la variable. -Determinando los posibles valores de la incógnita. -Calculando su probabilidad. -Graficando sus distribuciones. (MINEDUC, 2015, p. 124)

3° medio	3.Tomar decisiones en situaciones de incerteza que involucren el análisis de datos estadísticos con medidas de dispersión (MINEDUC, 2019a, p.110)
4° medio	2. Fundamentar decisiones en situaciones de incerteza, a partir del análisis crítico de datos estadísticos y con base en los modelos binomial y normal. (MINEDUC, 2019a, p.111)

Nota. Elaboración propia.

También en 4° medio son abordados algunos modelos de probabilidad, como se puede apreciar en las Tablas 3 y 4. Primero se estudia un modelo para variables aleatorias discretas, la distribución binomial, sobre el cual se propone incluir: su función de probabilidad y representación en gráfico de barras, sus parámetros (n: número de ensayos, p: probabilidad de éxito y q: probabilidad de fracaso), asignación de probabilidades mediante el enfoque clásico y frecuencial de probabilidad, y cálculo de su media, desviación estándar y varianza.

Tabla 4

Indicadores de evaluación sobre variable aleatoria y modelos de probabilidad en los Programas de Estudio

Curso escolar	Indicadores de evaluación
2° medio	-Reconocen la diferencia entre las variables utilizadas en álgebra y las variables aleatorias. -Definen variables aleatorias finitas en experimentos aleatorios. -Determinan los valores que puede tomar la variable aleatoria finita. -Aplican correctamente la terminología $X = x_i$, en la cual los x_i representan los valores que puede tomar la variable aleatoria. -Determinan las probabilidades de una variable aleatoria aplicando la terminología P(X = xi). -Elaboran tablas y gráficos para representar la distribución de una variable aleatoria finita. (MINEDUC, 2016, p.154)
3° medio	-Extraen e interpretan información estadística, calculando medidas de dispersión para comparar situaciones. -Analizan datos, calculando las medidas de dispersión para tomar decisiones. -Representan la información y utilizan las medidas de dispersión para comunicar alguna decisión. (MINEDUC, 2021a, p.36) -Calculan la desviación estándar de forma manual o utilizando herramientas digitales para interpretar datos de una situación. -Explican sus respuestas, utilizando datos extraídos del contexto y terminología relacionada con las medidas de dispersión. (MINEDUC, 2021a, p.58)
4° medio	-Interpretan datos de un experimento aleatorio dicotómico como la base del modelo binomial. -Comparan la probabilidad de una variable aleatoria y la frecuencia relativa de un suceso en un experimento aleatorio. -Evalúan las diferentes posibilidades en un experimento aleatorio y determinan su probabilidad. -Elaboran diagramas de árboles para representar las probabilidades de los diferentes sucesos de un experimento. (MINEDUC, 2021b, p.39) -Utilizan la tabla de probabilidades para determinar la probabilidad de tomar, en forma aleatoria, un dato de una población distribuida normalmente. -Evalúan la pertinencia de usar modelos binomial o normal para interpretar situaciones de incerteza. -Evalúan los alcances y límites de un argumento estadístico o probabilístico antes de tomar una decisión. (MINEDUC, 2021b, p.51)

	-Determinan la probabilidad de intervalos dentro de una distribución normal, utilizando la tabla probabilística para Z. -Determinan porcentajes de situaciones basándose en el cálculo de las probabilidades. -Evalúan la veracidad de proposiciones utilizando el concepto de distribución normal. - Determinan si los datos corresponden a una distribución normal. -Determinan el promedio y la desviación estándar. -Representan datos e información sobre el contexto utilizando histogramas. -Conjeturan sobre la tendencia de los datos utilizando un histograma. -Describen la tendencia de los datos utilizando la distribución normal o aproximaciones de ella. (MINEDUC, 2021b, p.66)

Nota. Elaboración propia.

Posteriormente es introducido un modelo para variable aleatorias continuas, la distribución normal, en el cual se fomenta el tratamiento de: su función de probabilidad y representación gráfica en una curva normal e histograma; sus parámetros (μ: media y σ: desviación típica); algunas propiedades de su función de probabilidad (es simétrica con respecto a la media y esta última es la mediana y la moda de la distribución) y la propiedad de los intervalos centrales ($\pm 3\sigma$); asignación de probabilidades por medio del enfoque frecuencial y clásico de probabilidad; la tipificación de la variable aleatoria con distribución normal; propiedades para calcular probabilidades con distribución normal estándar y cálculo de probabilidades empleando la tabla de probabilidades de la distribución normal estándar.

Finalmente se sugiere abordar la aproximación de la distribución binomial por la normal de manera gráfica, por lo que es propuesta una secuencia de distribuciones binomiales, representadas en gráficos de barras, donde varía su parámetro n y plantean preguntas para describir y conjeturar sobre la tendencia de los datos.

1.3.2 Normativas curriculares internacionales

Los estándares que regulan la educación matemática obligatoria en Estados Unidos de América, como los Principios y Estándares para la Educación Matemática (National Council of Teachers of Mathematics [NCTM], 2000) y los Estándares Estatales Básicos Comunes (National Governors Association Center for Best Practices y el Council of Chief State School Officers [NGACBP y CCSSO], 2010), establecen para la asignatura matemática las metas (conocimientos y habilidades) que se necesitan lograr al término de cada grado escolar (ver Tabla 5) y ofrecen experiencias de enseñanza-aprendizaje de primer nivel. En tanto que los Lineamientos para la Evaluación y Enseñanza en Educación Estadística, conocido como proyecto GAISE (Franklin et al., 2005), tienen como propósito lograr en el trascurso de la educación escolar (pre-kindergarten a grado 12) estudiantes alfabetizados estadísticamente.

Tabla 5

Correspondencia entre grados escolares estadounidenses y edad

Edad	Grado escolar
4-5 años	*Pre-kindergarten
5-6 años	Kindergarten
6-7 años	1 grado
7-8 años	2 grado
8-9 años	3 grado
9-10 años	4 grado
10-11 años	5 grado
11-12 años	6 grado
12-13 años	7 grado
13-14 años	8 grado
14-15 años	9 grado
15-16 años	10 grado
16-17 años	11 grado
17-18 años	12 grado

*Grado no incluido en los Estándares Estatales Básicos Comunes

Nota. Elaboración propia.

1.3.2.1 La variable aleatoria y modelos de probabilidad en los estándares americanos

Los Principios y Estándares para la Educación Matemática (NCTM, 2000) están constituidos por diez estándares abordados en cada grado escolar con distinto énfasis: número, álgebra, geometría, medida, análisis de datos y probabilidad, resolución de problemas, razonamiento y prueba, comunicación, conexión y representación. Entre los grados 9 y 12, en el estándar análisis de datos y probabilidad, se sugiere introducir la idea de variable aleatoria (ver Tabla 6), abarcando: el estudio de variables aleatorias de tipo discretas; asignación de probabilidades a sus valores desde el enfoque frecuencial y clásico de probabilidad; determinación de su función de probabilidad; construcción de gráfico de barras que muestre información sobre la frecuencia relativa o probabilidad de sus valores; y cálculo de su media. También es abordada la resolución de problemas mediante la distribución binomial, puntualmente cálculo de probabilidades desde el enfoque frecuencia y clásico.

Tabla 6

Expectativa de aprendizaje sobre variable aleatoria en los Principios y Estándares para la Educación Matemática

Etapa	Expectativas de aprendizaje
Grados 9 a 12	-Comprender los conceptos de espacio muestral y distribución de probabilidad y construir espacios muestrales y distribuciones en casos simples

	-Usar simulaciones para construir distribuciones de probabilidad empíricas -Calcular e interpretar el valor esperado de variables aleatorias en casos simples (NCTM, 2000, p.324) -Encontrar solución a un problema utilizando la distribución binomial (NCTM, 2000, p.333)

Nota. Elaboración propia.

Los Estándares Estatales Básicos Comunes para las Matemáticas (NGACBP y CCSSO, 2010) están estructurado por ejes temáticos según el grado escolar: en pre-kinder, se aborda conteo y cardinalidad, operaciones y pensamiento algebraico, número y operaciones en base diez, medición y datos, y geometría; para 1 y 2 grado, se presentan operaciones y pensamiento algebraico, número y operaciones en base diez, medición y datos, y geometría; desde grado 3 a 5, se incluyen las operaciones y pensamiento algebraico, número y operaciones en base diez, número y operaciones con fracciones, medición y datos, y geometría; en los grados 6 y 7, se abarcan las razones y relaciones proporcionales, el sistema numérico, expresiones y ecuaciones, geometría, y estadística y probabilidad; en grado 8, el sistema numérico, expresiones y ecuaciones, funciones, geometría, y estadística y probabilidad; a partir de grado 9 hasta grado 12 se consideran los números y sistemas numéricos, álgebra, funciones, modelamiento, geometría, y estadística y probabilidad.

Particularmente entre los grados 9 y 12, el eje estadística y probabilidad propone el dominio *usar la probabilidad para tomar decisiones,* donde se sugiere estudiar la variable aleatoria centrándose en las de tipo discretas (ver Tabla 7). Así es incluida: su definición en la resolución de problemas; identificación de su dominio y recorrido; asignación de probabilidad a sus valores mediante el enfoque clásico y frecuencial de probabilidad; determinación de su distribución de probabilidad (función de probabilidad) y representación de esta en gráfico de barras; y cálculo de su media. También el eje temático señalado presenta el dominio *interpretación de datos categóricos y cuantitativos*, momento en que recomienda trabajar la distribución normal (ver Tabla7), considerándose: su representación gráfica en una curva normal e histograma; sus parámetros (μ: media y σ: desviación típica); asignación de probabilidades a través del enfoque frecuencial y clásico de probabilidad; la tipificación de la variable aleatoria con distribución normal; propiedades para calcular probabilidades con distribución normal estándar y cálculo de probabilidades empleando la tabla de probabilidades de la distribución normal estándar.

Tabla 7

Dominios y estándares sobre variable aleatoria y modelos de probabilidad en los Estándares Estatales Básicos Comunes para las Matemáticas

Etapa	Dominios y Estándares
Grados 9 a 12	*Interpretación de datos categóricos y cuantitativos* -Resumir, representar e interpretar los datos en una única cuenta o variable de medición 1. Representar los datos con gráficos sobre la recta real (histogramas) 4. Usar la media y la desviación estándar de un conjunto de datos para ajustarlos a una distribución normal y estimar porcentajes de población. Reconocer que hay conjuntos de datos para los que tal procedimiento no es apropiado. Utilizar calculadoras, hoja de cálculo y tablas para estimar áreas bajo la curva normal. (NGACBP y CCSSO, 2010, p.81) *Usar la probabilidad para tomar decisiones* -Calcular el valor esperado y utilizarlo para resolver problemas 1. Definir una variable aleatoria para una cantidad de interés, asignando un valor numérico a cada suceso de un espacio muestral. Graficar la distribución de probabilidad correspondiente utilizando el mismo gráfico como para distribuciones de datos 2. Calcular el valor esperado de una variable aleatoria e interpretarlo como la media de la distribución de probabilidad. 3. Desarrollar una distribución de probabilidad para una variable aleatoria definida en un espacio muestral en el que se puede calcular las probabilidades teóricas. Encontrar el valor esperado 4. Desarrollar una distribución de probabilidad para una variable aleatoria definida en un espacio muestral en el que las probabilidades se asignan empíricamente. Encontrar el valor esperado -Usar la probabilidad para evaluar los resultados de las decisiones 5. Examinar los posibles resultados de una decisión asignando probabilidades a los valores de rentabilidad y encontrando los valores esperados: a) encontrar el pago esperado para un juego de azar; b) evaluar y comparar estrategias sobre la base de los valores esperados. (NGACBP y CCSSO, 2010, p.82-83)

Nota. Elaboración propia.

1.3.2.2 La variable aleatoria y modelos de probabilidad en el proyecto GAISE

Los Lineamientos para la Evaluación y Enseñanza en Educación Estadística (Bargagliotti et al., 2020; Franklin et al., 2005) están estructuradas en tres niveles crecientes de complejidad (A, B y C), para cada uno de ellos son propuestos seis componentes, cuatro componentes de proceso para resolver problemas (formular preguntas; recolectar datos; analizar datos; interpretar resultados), una sobre la naturaleza de la variabilidad y otra enfocada en la variabilidad. Además, se incorpora un apartado relativo al rol de la probabilidad en estadística solo para los niveles A y C. En este último es propuesto abordar la variable aleatoria (ver Tabla 8), principalmente la de tipo discreta, abarcando su definición en problemas, representación de su distribución de probabilidad en tabla, y cálculo de su media y desviación estándar.

También en el contexto señalado se considera el tratamiento de la distribución normal (ver Tabla 8), incluyendo sus parámetros (μ: media y σ: desviación típica), su representación gráfica en una curva normal e histograma, su propiedad de los intervalos centrales ($\pm 3\sigma$), la asignación de probabilidades por medio del enfoque frecuencial y clásico de probabilidad, y la tipificación de la variable aleatoria con distribución normal.

Tabla 8

La variable aleatoria y distribución normal en el proyecto GAISE

Nivel	Rol de la probabilidad
Nivel C	-Los conceptos de probabilidad que se necesitan para la estadística introductoria (con énfasis en análisis de datos) incluyen interpretaciones de frecuencia relativa de los datos, distribuciones de probabilidad como modelos de poblaciones de medidas, una introducción a la distribución normal como un modelo para la distribución muestral, e ideas básicas de valor esperado y variables aleatorias (Franklin et al., 2005, p.87) -Además, como muchos de los procedimientos de inferencia estándar están basados en la distribución normal, los estudiantes deben estar en condiciones de evaluar probabilidades usando la distribución normal, preferiblemente con ayuda de la tecnología (software y calculadora gráfica). (Franklin et al., 2005, p.88) -La distribución normal tiene la siguiente característica descrita por la regla empírica: i) se espera que aproximadamente el 68% de las observaciones se encuentren dentro de una desviación estándar de la media; ii) aproximadamente el 95% de las observaciones caigan dentro de dos desviaciones estándar de la media; iii) todas o casi todas las observaciones estén dentro de tres desviaciones estándar de la media (Bargagliotti et al., 2020, p.79)

Nota. Elaboración propia.

1.4 Marco teórico

La presente investigación se fundamenta en herramientas del Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos (EOS), el cual ha sido desarrollado por Godino et al. (2007; 2020). Esta teoría posee diversas herramientas que hacen posible realizar una investigación desde distintas perspectivas (estudiantes, profesor, currículo o enseñanza), relacionándolas e integrándolas. A continuación, son descritas aquellas herramientas teóricas utilizadas.

1.4.1 Prácticas matemáticas, objeto matemático y significado

El EOS posee su origen en las ideas de situación-problema (actividad, problema, ejercicio, etc.) y prácticas matemáticas realizadas para su resolución. Aquellas prácticas corresponden a acciones o expresiones desarrolladas con el propósito de hallar la solución a un problema, informar su solución o generalizarlo (Godino y Batanero, 1994). Entre estas se distinguen (Godino, 2018): las prácticas personales, llevadas a cabo por un individuo; y las prácticas institucionales, desarrolladas de manera conjunta en una institución que comparte los mismos instrumentos, funcionamientos y reglas. También en las prácticas matemáticas se diferencian las prácticas operativas, realizadas al aplicar ciertos objetos matemáticos (procedimientos, propiedades, conceptos y lenguaje) mientras se resuelve un problema, y las prácticas discursivas, relativas a los argumentos y justificaciones que fundamentan su solución o se emplean en la explicación de alguna propiedad.

Para el Enfoque Ontosemiótico los objetos se originan de las prácticas matemáticas, y corresponden a toda entidad real o imaginaria que participa en la práctica matemática y que puede individualizarse (Godino et al., 2007). En tanto que el significado de un objeto es una entidad integrada por (Godino, 2018): el nombre del objeto; el objeto institucional, constituye el conocimiento objetivo y depende de la institución de referencia; el objeto personal, correspondiente al conocimiento subjetivo; y las prácticas vinculadas a la resolución de situaciones-problemas del que surge el aquel conocimiento objetivo. Desde este punto de vista, es posible comprender la formación y desarrollo de los conocimientos, tanto mental como culturalmente.

1.4.2 Sistema de prácticas

Godino et al. (2007) proponen que el sistema de prácticas llevadas a cabo dentro de una institución con el propósito de dar solución a una situación-problema se denomina significado institucional de un objeto matemática, y aquellas atribuidas a una persona se llama significado personal. Esta diferenciación, entre el significado institucional y personal es importante tanto para la descripción y explicación de las interacciones profesor-alumno en el proceso enseñanza-aprendizaje, como para la indagación de la actividad matemática.

Dentro del significado institucional de un objeto matemático es posible distinguir cuatro tipos (ver Figura 1):

- El significado institucional de referencia, es aquel que el investigador o profesor utiliza como pauta para elaborar el significado a trabajar en el aula o llevar a cabo un estudio. Para establecerlo son empleados libros de textos científicos o lineamientos curriculares.
- El significado institucional pretendido, corresponde a las prácticas en torno a un objeto matemático integradas en la planificación de la enseñanza. Para determinarlo son utilizados libros de texto o diseño de un proceso de estudio.
- El significado institucional implementado, hace referencia al sistema de prácticas aplicadas de manera efectiva en el aula por el investigador o profesor, y orienta la elaboración de

evaluaciones que posteriormente deben responder los estudiantes, dentro del proceso enseñanza-aprendizaje.

- El significado institucional evaluado, es el conjunto de prácticas empleado por el profesor o investigador, por medio de situaciones-problemas o pautas de observación, para evaluar los aprendizajes sobre un objeto matemático.

Figura 1

Tipos de significados institucionales y personales

Nota. Extraída de Godino et al. (2008, p.6).

En el significado personal de un objeto matemático se diferencian tres tipos (Godino et al., 2007):

- El significado personal global, corresponde al sistema de prácticas que potencialmente un individuo, como el estudiante, podría lograr sobre un objeto matemático específico.
- El significado personal declarado, hace referencia a las prácticas plasmadas por el estudiante en alguna evaluación presentada por el investigador o profesor, las cuales puede ser adecuadas o erróneas.
- El significado personal logrado, son las prácticas realizadas por el estudiante en torno a un objeto matemático concreto y coherentes con las propuestas en el significado institucional, aquellas que no son congruentes con las propuestas por la institución se denominan errores de aprendizaje.

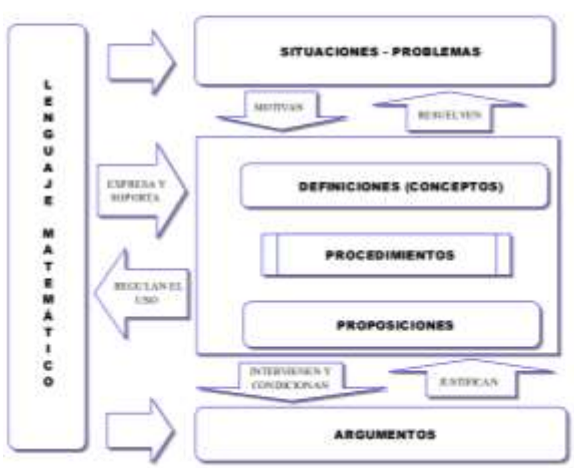
1.4.3 Tipología de objetos matemáticos primarios

Según el EOS en el desarrollo y evaluación de una práctica matemática intervienen diversos objetos matemáticos primarios (ver Figura 2) que se pueden clasificar en la siguiente tipología (Godino, 2002; Godino 2008):

- Situaciones-problemas, son las tareas, actividades, problemas o ejercicios asociados a un objeto matemático y que promueven la actividad matemática.
- Lenguaje, corresponde a los términos, notaciones, expresiones o gráficos utilizados para enunciar una situación-problema, resolverla, generalizar o describir su solución.
- Conceptos, descripciones o definiciones relacionadas a un objeto matemático, que el individuo conoce previamente y aplica en la resolución de una situación-problema.
- Proposiciones, hacen referencia a las características o propiedades de un objeto matemático utilizadas para solucionar una situación-problema, además cada atributo de un objeto lo vincula a otros distintos y ayuda a la evolución del significado del objeto de interés.
- Procedimientos, son los algoritmos, técnicas de cálculos u operaciones desarrolladas para resolver una situación-problema.
- Argumentos, correspondiente al razonamiento utilizado en la comprobación, explicación y justificación de la solución a una situación-problema, o también entendidos como los enunciados empleados para aprobar o justificar proposiciones, procedimientos y/o la solución.

Figura 2

Configuración de objetos matemáticos primarios



Nota. Extraída de Godino et al. (2008, p.7).

Así, “*un conjunto de situaciones-problemas vinculadas recíprocamente, que comparten sus representaciones, procesos o soluciones similares constituyen un campo de problema*" (Godino, 1999, p.199). Además, las redes que se establecen de la interacción entre los objetos matemáticos primarios se llaman configuración, esta puede ser: cognitiva, propia de un individuo, y posibilita un análisis específico y una mejor comprensión del aprendizaje de los estudiantes (Godino, 2012); epistémica, específica de una institución, y es una herramienta útil para analizar textos matemáticos de diferentes épocas (Font y Godino, 2006).

Godino et al. (2020) sostienen que la identificación explicita de los objetos matemáticos primarios que participan y surgen en las prácticas matemáticas hacen posible predecir potenciales conflictos semióticos, evaluar la comprensión de estudiantes y reconocer aquellos objetos que necesitan ser institucionalizados en un momento específico durante el proceso de enseñanza-aprendizaje.

1.4.4 Comprensión y conflicto semiótico

El Enfoque Ontosemiótico (Godino et al., 2020) entiende la comprensión de un objeto matemático como una competencia, es decir, un individuo comprende un objeto específico cuando puede utilizarlo de forma idónea en diversas prácticas. En esta perspectiva, la comprensión de un individuo (Godino, 2002) se deduce de la observación de sus respuestas a situaciones-problemas, por lo que la evaluación de su comprensión se entiende como la correspondencia entre su significado personal y el significado institucional de un objeto matemático. La comprensión lograda por este en un momento específico escasamente puede ser nula o total, sino que incluirá algunos elementos de la diversidad de los objetos primarios involucrados.

Otra noción importante para el EOS es la de conflicto semiótico (Godino et al., 2007), correspondiente a un desajuste entre los significados asignados por dos entidades (individuo y/o institución) a una expresión. Existen distintos tipos según su origen (Godino, 2008): epistémico, si la discordancia surge entre significados institucionales; cognitivo, si la disparidad emerge entre las prácticas matemáticas que constituyen el significado personal; interaccional, si el desajuste se produce entre las prácticas desarrolladas por estudiante-estudiante o profesor-estudiante.

1.5 Objetivos de la investigación

Esta investigación tiene como propósito los siguientes objetivos principales:

Objetivo 1 (O1): *Desarrollar una revisión bibliográfica sobre la enseñanza y aprendizaje de variable aleatoria y distribuciones binomial y normal.*

El presente objetivo se fundamenta, primero en indagar investigaciones centradas en estudiar el origen y desarrollo histórico de variable aleatoria, distribución binomial y/o distribución normal, y estudios que analizan su tratamiento en libros de estadística y probabilidad o textos escolares. Además de explorar investigaciones sobre propuestas de enseñanza de los temas en cuestión, tanto a nivel escolar como universitario. Segundo, en indagar estudios sobre el aprendizaje de variable aleatoria, distribución binomial o distribución normal, en el contexto escolar e inicios de la universidad, identificando dificultades y sesgos cognitivos que los estudiantes pueden manifestar durante este proceso.

Objetivo 2 (O2): *Analizar el tratamiento de variable aleatoria y distribuciones binomial y normal en la normativa curricular y libros de texto chilenos.*

El interés de este objetivo radica, inicialmente en identificar en las Bases Curriculares y Programas de Estudio de matemática las situaciones-problemas, lenguaje, proposiciones, procedimientos y argumentos sobre variable aleatoria, distribución binomial y distribución normal, así establecer el significado institucional de referencia de aquellos temas según el currículo escolar chileno. Posteriormente reconocer en libros de texto chilenos la red de objetos matemáticos primarios en torno a los temas en cuestión, de esta manera determinar la configuración epistémica que define el significado institucional pretendido de variable aleatoria y distribuciones binomial y normal en los libros analizados.

Objetivo 3 (O3): *Construir y validar un instrumento orientado a evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en el contexto escolar chileno.*

Los estudios en torno a la comprensión de los temas en cuestión son prácticamente inexistentes en el contexto chileno. Por tanto, es de interés, primero establecer los campos de problemas y situaciones-problemas sobre variable aleatoria y distribuciones binomial y normal que estructuran el instrumento, mediante la indagación de estudios epistemológicos previos y el análisis de normativas curriculares chilena e internacionales. Luego seleccionar un conjunto inicial de ítems representantes de aquellas situaciones-problemas, a partir de la indagación de investigaciones previas y análisis de libros de texto. Después realizar la validez de contenido de los ítems por medio de un juicio de expertos, donde su evaluación permita obtener un conjunto final de ítems.

Posteriormente, es de interés desarrollar una prueba piloto de la versión inicial del instrumento y llevar a cabo su validez de constructo, donde mediante un análisis factorial exploratorio se indague su estructura y por medio de un análisis factorial confirmatorio se realice una aproximación a confirmar aquella estructura. También analizar la fiabilidad de la versión final instrumento, el índice de dificultad y discriminación de los ítems que lo componen. Finalmente determinar el significado institucional evaluado en el instrumento sobre variable aleatoria y distribuciones binomial y normal.

Objetivo 4 (O4): *Evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en estudiantes egresados de educación escolar chilena.*

El interés de este objetivo se basa, en aplicar la versión final del cuestionario a estudiantes egresados de educación escolar chilena, con la finalidad de valorar la comprensión de los participantes sobre los temas en cuestión en términos del EOS. Esto significa, identificar el lenguaje, conceptos, proposiciones, procesos y argumentos desarrollados por los participantes para responder las situaciones-problemas que componen el cuestionario. Así establecer el significado personal logrado por los participantes en torno a variable aleatoria y distribuciones binomial y normal, y analizar la correspondencia con su significado institucional evaluado en el cuestionario. Además de determinar los conflictos semióticos que manifiesten los estudiantes.

1.6 Hipótesis de la investigación

De acuerdo con los resultados de estudios previos sobre los temas de interés, son plantadas las hipótesis iniciales de la presente investigación en torno a las expectativas acerca de los objetivos expuestos anteriormente.

Hipótesis 1 (H1): *El análisis de investigaciones previas sobre la enseñanza y aprendizaje de variable aleatoria y/o distribuciones binomial y normal muestra que han sido temas limitadamente investigados*

La presente hipótesis se sustenta en el estudio de antecedentes realizado por: i) Tauber (2001), la cual evidenció que no existen investigaciones en torno a la enseñanza y aprendizaje de distribución normal, aunque hay aislados estudios que valoran el rendimiento en cursos de introducción a la estadística y abordan la normal como una componente de la investigación no como tema central; ii) Alvarado (2007), quien observó que hay escasas investigaciones sobre la enseñanza del teorema central del límite, y en particular en torno a la aproximación de la distribución binomial por la norma para valores grandes de n (número de ensayos); y iii) Ruiz (2013), la cual evidenció que existen pocas investigaciones sobre variable aleatoria desde la perspectiva matemática y en torno a su aprendizaje, debido a que la mayoría de los estudios se centran en otros temas que se vinculan a la variable aleatoria, como por ejemplo la aleatoriedad.

Hipótesis 2 (H2): *En el tratamiento de variable aleatoria y distribuciones binomial y normal existen diferencias entre el sugerido por la normativa curricular chilena y propuesto por los libros de texto chilenos.*

Esta hipótesis se fundamenta, en la disparidad de resultados obtenidos en cada investigación relativa a: i) la indagación de la normativa curricular española y el análisis de libros de texto españoles sobre variable aleatoria (Ortiz, 2002) y distribución normal (Valverde, 2007); ii) el análisis de textos escolares franceses en torno a la variable aleatoria (Doukhan y Gueudet, 2019); iii) la indagación de libros de texto chinos, japoneses, singapurenses, estadounidenses y británicos sobre distribución binomial (Li et al., 2021); y) el análisis de textos escolares indonesios en torno a la variable aleatoria y distribuciones binomial y normal (Setiawan, 2020).

Hipótesis 3 (H3): *El instrumento diseñado para evaluar la comprensión de variable aleatoria y distribuciones binomial y normal posee buenas características psicométricas*.

La presente hipótesis se apoya, en que la construcción del cuestionario que valora los temas de interés se basa en la metodología empleada por Batanero y Diaz (2005; 2006) y por Gómez-Torres et al. (2016), quienes han obtenido instrumentos con buenas características psicométricas. Aquella metodología consiste en: i) a partir del análisis de contenido de documentos que son referencia para la enseñanza (normativas curriculares o textos de estadística) e indagación de investigaciones previas, son establecidas las variables (especificaciones) que componen el instrumento de evaluación; ii) desde investigaciones previas son elegidos dos a tres ítems para cada especificación y mediante juicio de expertos es seleccionado aquel ítem idóneo que integran el instrumento; y iii) se desarrolla una prueba piloto del instrumento, donde se indaga su validez de constructo (mediante un análisis factorial) y su fiabilidad, además de analizar el índices de dificultad y discriminación de los ítems que lo componen.

Hipótesis 4 (H4): *Los participantes tienen una comprensión insuficiente de variable aleatoria y/o distribuciones binomial y normal*.

Aquella hipótesis se sustenta, en los resultados conseguidos en trabajos previos sobre evaluación de la comprensión de los temas de interés en estudiantes universitarios. Puntualmente dificultades manifestadas por los participantes, relacionadas a la variable aleatoria (Ruiz y Albert, 2005), distribución binomial (Toledo et al., 2019; Vilca, 2015), distribución normal (Tauber et al., 2004; Yuan y Li, 2012) y aproximación de la binomial por la normal (Alvarado y Batanero, 2007).

1.7 Organización de la investigación y resumen de la metodología

La presente investigación está organizada en seis estudios complementarios (cuatro publicados y uno en evaluación), vinculados a los objetivos expuestos previamente, y que posibilitan el logro de cada uno de ellos. A continuación, se expone de manera breve el método empleado en cada uno de ellos, agrupados según el objetivo que lo motiva.

Para el logro del *O1* se lleva a cabo el estudio publicado en el Boletín de Estadística e Investigación Operativa (BEIO) (Bizet et al., 2022), donde se realiza una revisión bibliográfica sobre la enseñanza- aprendizaje de los temas en cuestión (ver Capítulo 2). Este estudio se desarrolla desde un enfoque cualitativo y es de tipo descriptivo. Se utiliza el diseño de investigación documental, mediante el cual se realiza una revisión de la literatura en bases de datos bibliográficas, un servidor de internet, y revistas y actas de congreso de educación matemática y estadística. El proceso de búsqueda se lleva a cabo en idioma español e inglés, con límite de año de publicación (1990 a 2020) y empleando descriptores específicos (variable aleatoria y educación matemática; variable aleatoria y educación estadística; variable aleatoria y comprensión; distribución binomial y educación estadística; distribución normal y educación estadística). Así se obtiene una muestra de 52 trabajos en los que se analizan elementos como: enfoque para abordar el problema de investigación, muestra, contexto, instrumento, resultados y conclusiones.

Respecto al *O2*, para su cumplimiento se desarrollan dos estudios sobre la indagación a la normativa curricular y libros de texto chilenos en torno a los temas de interés, aquellos son expuestos en el Capítulo 3. El primer estudio se encuentra publicado en Profesorado, revista de currículum y formación de profesorado (Bizet et al., 2023a) y el segundo está publicado en PNA, revista de investigación en Didáctica de la Matemática (Bizet et al., 2023b). Cada estudio es realizado bajo el enfoque cualitativo y es de tipo exploratorio-descriptivo. Se emplea el diseño de investigación documental y a través de la técnica de análisis de contenido son analizadas principalmente las Bases Curriculares para los grados 9 a 12 (14 a 18 años) y 5 libros de texto chilenos de matemática dirigido a los grados 10 a 12 (15 a 18 años). El estudio abarca las etapas de indagar los capítulos sobre variable aleatoria, distribución binomial y distribución normal, establecer los objetos matemáticos primarios en torno a estos temas y elaborar tablas que resumen la información obtenida.

Sobre el *O3*, para su logro son realizados dos estudios presentados en el Capítulo 4. El primero se encuentra publicado en la revista Educación Matemática (Bizet et al., 2023c) y aborda el proceso de construcción y validez de contenido de un cuestionario para evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en el contexto escolar chileno. Este es un estudio desarrollado desde el enfoque mixto y es de tipo exploratorio-descriptivo, se usa un diseño transversal y organiza en tres fases:

i. *Construcción de una guía de situaciones-problemas*. En esta fase, desde el análisis de investigaciones previas, tanto sobre el origen y desarrollo histórico de la variable aleatoria, distribución binomial y distribución normal, como en torno a su tratamiento en libros de estadística, son establecidos los campos de problemas de estos temas. Posteriormente, por medio de la técnica de análisis de contenido, son indagadas las Bases Curriculares y normativas curriculares internacionales, y definidas unidades de análisis, las cuales son clasificadas en los campos de problemas previamente establecidos. Finalmente, las unidades de análisis son comparadas y reducidas, y a partir de estas se infieren situaciones-problemas sobre variable aleatoria, distribución binomial y distribución normal.
ii. *Selección de un conjunto inicial de ítems a partir de la literatura*. En la segunda fase, para cada situación-problema que componen la guía se elige desde la literatura un trío de ítems que la representen (actividades, tareas o ejercicios concretos).
iii. *Selección de un conjunto final de ítems a partir de la validez de contenido por juicio de expertos*. En la última fase, participan seis doctores en educación matemática o psicología,

quienes utilizando una escala Likert valoran el grado en que cada ítem es adecuado para evaluar la situación-problema donde fue clasificado, además eligen del trío aquel ítem que consideran que evalúa de mejor manera la situación-problema. Después se desarrolla un análisis descriptivo de los resultados por medio de indicadores estadísticos sobre el centro y la dispersión de las valoraciones de los expertos, donde del trío de ítems propuesto para cada situación-problema es seleccionado aquel con media más alta.

El segundo estudio que permite el logro del *O3*, está en proceso de evaluación en una revista sobre investigaciones en educación (Bizet et al., en evaluación) y presenta la validez de constructo y fiabilidad del instrumento construido previamente. Aquel estudio es desarrollado desde un enfoque cuantitativo y es de tipo descriptivo-correlacional. Se utiliza el diseño de investigación transversal e implementa el cuestionario en una muestra de 80 estudiantes universitarios chilenos que en la universidad no han cursado alguna asignatura sobre temas estocásticos. La validez del instrumento es indagada mediante un análisis factorial exploratorio y confirmatorio, y su fiabilidad empleando el coeficiente de alfa de Cronbach.

Además, el Capítulo 4 abarca la indagación del índice de dificultad y discriminación de cada ítem que compone la versión final del instrumento, y la identificación del significado institucional de los temas de interés evaluado en aquel cuestionario.

Para el cumplimiento del *O4*, se lleva a cabo el estudio expuesto en el Capítulo 5, el cual aborda la evaluación de la comprensión de variable aleatoria y distribuciones binomial y normal en estudiantes egresados de educación escolar chilena. El estudio se realizar bajo una perspectiva mixta y es de tipo exploratorio-descriptivo. Se emplea el diseño no experimental (transversal) y aplica el cuestionario en una muestra de 101 estudiantes universitarios chilenos que en la universidad no han cursado alguna asignatura sobre temas estocásticos.

Primero se desarrolla un análisis cualitativo de las respuestas, donde a través de la técnica de análisis de contenido y mediante un proceso cíclico e inductivo son definidas y descritas categorías en cada ítem. En las categorías de respuestas parciales y correctas de cada ítem, son identificados los objetos matemáticos primarios movilizados por los participantes para su solución y su frecuencia. Además, se identifican conflictos semióticos presentes principalmente en las categorías de respuestas erróneas y parcialmente correctas. Posteriormente, en el análisis cuantitativo, se realiza un análisis estadístico descriptivo para cada uno de los ítems, considerando la puntuación media obtenida por los participantes en cada ítem, la dispersión de aquellas puntuaciones, la puntuación total en el cuestionario y su media. Además, se incluye la puntuación total diferenciada por grupo que integran la muestra.

CAPÍTULO 2: ANTECENTES

2.1 Introducción

Este capítulo expone una revisión bibliográfica que fundamenta los demás estudios incluidos en la presente tesis doctoral. Aquella es abordada en el apartado 2.2 mediante el Estudio 1. What does the research tell us about the understanding of the random variables and its probability distributions? publicado en la revista BEIO. Primero son indagadas las investigaciones más relevantes en torno a la enseñanza, en el ámbito escolar y universitario, de la variable aleatoria y distribuciones binomial y normal, donde se abordan sus aspectos epistemológicos y didácticos. Posteriormente son analizadas las investigaciones más importantes sobre el aprendizaje, en el contexto escolar y universitario, de los temas en cuestión, allí se tratan sus aspectos cognitivos.

En el apartado 2.3 de este capítulo, se continua con el desarrollo de la revisión bibliográfica, incorporándose investigaciones excluidas en la publicación de BEIO, debido a que en la búsqueda de la literatura se integraron nuevos descriptores y extendió el límite de año de publicación. Esta información se presenta organizada de la manera previamente señalada.

2.2 Estudio 1[1]. What does the research tell us about the understanding of the random variables and its probability distributions?

Bizet, V., Molina-Portillo, E., Ruz, F. y Contreras, J.M. (2022). What does the research tell us about the understanding of the random variables and its probability distribution? Boletín de Estadística e Investigación Operativa, 38(3), 1-22.

Abstract

The random variable represents one of the key concepts in the modeling of random phenomena through probability distributions. Therefore, the objective of this study is to analyze and describe the main investigations that the literature reports on the random variable and its probability distribution, through a bibliographic review. The results show the existence of some teaching proposals around this notion, which are characterized by the use of technology. In addition, various cognitive difficulties and biases are identified during the learning of the random variable. In this way, it is hoped that these results will contribute to strengthening the instructional design process carried out by teachers, in order to help students understand the random variable and its probability distribution.

Keywords: random variable, probability distribution, understanding, statistics education.

MSC Subject classifications: 97K50, 97K60.

2.2.1 Introducción

El conocimiento estocástico se ha destacado como un aspecto fundamental para entender el mundo que nos rodea, pues entrega herramientas para razonar en situaciones de incertidumbre. En este contexto, la comprensión del objeto variable aleatoria representa uno de los conceptos fundamentales en la modelización de fenómenos aleatorios por medio de distribuciones de probabilidad, además es considerada desde hace más de 40 años como un concepto fundamental de la estocástica (Heitele, 1975).

[1] Para evitar confusión y repetición en la numeración, se ha modificado la relativa a los epígrafes pertenecientes a los estudios que forman parte del compendio, haciendo referencia al capítulo en que se ubica cada estudio.

El estudio de la variable aleatoria se aborda tanto en los últimos años de educación obligatoria dentro de la asignatura de matemática, como en los cursos universitarios de introducción a la estadística. En el ámbito escolar se ha establecido dentro de diversas directrices curriculares internacionales, entonces para ilustrar esta situación de manera uniforme, en el presente estudio se ha adoptado la denominación K-12 (Franklin et al., 2005) para organizar los doce grados de educación obligatoria. Por ejemplo, en los Principles and standards for school mathematics (NCTM, 2000), entre los grados 9 y 12 (14 a 18 años) se incluye la comprensión de su distribución de probabilidad y el cálculo e interpretación de su esperanza matemática. Mientras que el currículo español (MECD, 2015) se inicia su estudio en bachillerato, entre los grados 11 y 12 (16 a 17 años) y el eje temático Estadística y Probabilidad, por medio de los siguientes contenidos: variables aleatorias discretas y continuas, distribuciones de probabilidad como la binomial y la normal, tipificación de la distribución normal, cálculo de probabilidades mediante la aproximación de la distribución binomial por la normal, entre otros.

Por su parte, la Didáctica de la Matemática se ha interesado en estudiar la enseñanza y aprendizaje de la variable aleatoria, logrando reconocer diversas dificultades en torno a la comprensión de conceptos vinculados a esta y su aplicación en el estudio de tópicos estocásticos posteriores, en los distintos niveles educativos. De esta forma se ha identificado un problema vigente en la comprensión de la variable aleatoria y su distribución de probabilidad, por ello el objetivo de este estudio es analizar de manera descriptiva las investigaciones que reporta la literatura sobre estos conceptos a través de una revisión bibliográfica.

En lo que sigue, este trabajo se organiza presentando el método utilizado (Sección 2.2.2). Luego en la Sección 2.2.3, son dados a conocer los principales resultados del análisis de las investigaciones seleccionadas, distribuidos en dos subapartados: Sección 2.2.3.1, enfocada en la enseñanza de la variable aleatoria y su distribución de probabilidad, identificando los elementos epistemológicos y didácticos de los conceptos; Sección 2.2.3.2, centrada en el aprendizaje de la variable aleatoria y su distribución de probabilidad, reportando aspectos cognitivos como los niveles de razonamiento en torno a los conceptos, algunas dificultades y sesgos del razonamiento probabilístico presentados por los estudiantes. Finalmente, en la Sección 2.2.4, se exponen las conclusiones e implicaciones para la enseñanza y aprendizaje de los conceptos en cuestión.

De esta forma, se proyectan los resultados de este trabajo como un insumo valioso para aportar y apoyar al profesorado en el diseño instruccional de experiencias de enseñanza de la variable aleatoria y su distribución de probabilidad.

2.2.2 Método

En este trabajo se llevó a cabo una revisión de la literatura (Hernández et al., 2014) consultando en bases de datos bibliográficas (Web of Science, Scopus, Math Educ y Proquest), un servidor en internet (Google Scholar), revistas y actas de congresos de educación matemática y estadística.

El proceso de búsqueda se realizó tanto en idioma español como inglés, con límite de año de publicación (1990 a 2020) y a través de los siguientes descriptores: variable aleatoria y educación matemática; variable aleatoria y educación estadística; variable aleatoria y comprensión; distribución binomial y educación estadística; distribución normal y educación estadística.

Las investigaciones fueron seleccionadas según su relación estrecha con el objeto de estudio, sin interés en el método de investigación y su análisis se llevó a cabo considerando los siguientes elementos: enfoque para abordar el problema de investigación, muestra, contexto, instrumento, resultados y conclusiones. Así han sido elegidos 52 trabajos para su análisis descriptivo y la presentación de la información analizada se organizó en dos apartados, el primero presenta estudios relativos a la enseñanza de la variable aleatoria y su distribución de probabilidad, mientras que el segundo expone investigaciones sobre su aprendizaje

2.2.3 Resultados

2.2.3.1 Enseñanza de la variable aleatoria y su distribución de probabilidad

En esta sección se exponen estudios que abordan tanto aspectos epistemológicos (relativos al contenido) como didácticos (elementos que intervienen en el proceso de su enseñanza) de la variable aleatoria y su distribución de probabilidad.

Investigaciones sobre aspectos epistemológicos de la variable aleatoria y distribuciones de probabilidad

Los estudios llevados a cabo desde el ámbito epistemológico se caracterizan por indagar el conocimiento probabilístico en torno a la variable aleatoria y su distribución de probabilidad, abarcando su origen, desarrollo histórico y el estudio del concepto desde la matemática (Dinges, 2005; Ortiz, 2002; Ruiz, 2013; Tauber, 2001; Valverde, 2017). El trabajo de Dinges (2005) se centra en un estudio del origen y desarrollo histórico del concepto de variable, este contempla desde los inicios del análisis matemático, abarcando lo llevado a cabo por Leibniz (creador del concepto de función), Cantor (descubridor de la teoría de conjuntos), hasta la actualidad. Entre sus resultados destaca que la variable posee distintos significados, uno vinculado al concepto algebraico (ámbito determinista) y otro relacionado con procesos aleatorios (ámbito estocástico), aunque en ambos contextos el concepto de función es fundamental. Además, presenta una propuesta de enseñanza de la variable aleatoria para nivel universitario, desde el enfoque axiomático de la probabilidad, constituida por cuatro axiomas.

Mientras que Ortiz (2002) caracteriza el significado institucional de la variable aleatoria a partir del análisis de libros de texto de bachillerato, basándose en los tipos de elementos primarios de un objeto matemático propuesto por el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos (EOS) (Godino et al., 2007). Así por medio de la técnica de análisis de contenido (Fox, 1981) indaga 11 libros de texto de 1° de bachillerato en España (grado 11, 16 a 17 años) e identifica la presencia o ausencia de algunos objetos primarios de la variable aleatoria (situaciones-problemas, conceptos-definiciones y proposiciones-propiedades). Entre sus resultados destaca que solo en algunos libros se incluyen situaciones-problemas que involucran de manera explícita a la variable aleatoria, ya que en tres libros no aparecen actividades relacionadas con ella y solo en un caso existe mayor diversidad de situaciones que promuevan su resolución de problemas. Sobre las definiciones y propiedades, se refleja la ausencia del concepto en la mayoría de los textos analizados, mientras que en aquellos en los que se aborda, se realiza de forma superficial, centrándose únicamente en elementos como su distribución de probabilidad y/o su esperanza matemática.

Luego Ruiz (2014) desarrolla un estudio epistemológico de la variable aleatoria desde la disciplina matemática, fundamentado en la teoría EOS. De esta manera establece los distintos elementos primarios de la variable aleatoria (situaciones-problemas, lenguaje, conceptos-definiciones, proposiciones-propiedades, procedimientos y argumentos) para cada uno de los significados históricos de la probabilidad (intuitiva, clásica, frecuencial, subjetiva y axiomática), obteniendo como resultado cuatro configuraciones epistémicas (colección de objetos que emergen de ciertos tipos de prácticas matemáticas) en torno a la variable aleatoria. La autora a modo de conclusión propone recomendaciones para su enseñanza escolar, como la necesidad de mantener al estudiante en constante contacto con el concepto, inicialmente emplearlo en situaciones concretas y de forma intuitiva, para posteriormente usarlo explícitamente de manera informal. Además, identifica posibles dificultades en: comprender la noción de distribución; vincular el significado frecuencial y clásico de la probabilidad; relacionar los conceptos variables aleatoria y estadística; y comprender la aproximación de distribuciones de probabilidad de variables aleatorias de tipo discretas a continuas.

Por su parte Tauber (2001) se centra en indagar la distribución normal en libros de introducción a la estadística dirigido a carreras de ciencias sociales, basándose en herramientas del EOS. Por medio del análisis de contenido revisa 11 libros y obtiene como resultado el significado institucional de referencia de la distribución normal, en términos de sus elementos primarios (situaciones-problemas, lenguaje, conceptos-definiciones, proposiciones-propiedades, procedimientos y argumentos). La autora concluye afirmando que la información obtenida es valiosa para los profesores, quienes deben elaborar instrumentos de evaluación y/o unidades didácticas sobre el tópico.

A diferencia de Valverde (2017) quien se interesa en analizar en libros de texto de bachillerato la presencia de la distribución normal, considerando como fundamento teórico el EOS. Utiliza como muestras dos libros de texto españoles con su respectivo manual de ejercicios y los examina mediante el método de análisis de contenido. Como resultado obtiene el significado institucional de referencia de la distribución normal para nivel escolar, en términos de sus elementos primarios (situaciones-problemas, lenguaje, conceptos-definiciones, proposiciones-propiedades, procedimientos y argumentos). Así constata que una gran cantidad de definiciones expuestas en los libros son confusas, utilizan un lenguaje formal distante al nivel escolar y existen escasos ejemplos. Además, evidenció un alto grado de idoneidad epistémica debido a que una gran cantidad de objetos primarios identificados por Tauber (2001) se observaron en los libros de texto escolares.

Investigaciones sobre la didáctica de la variable aleatoria

Aquellas investigaciones interesadas en aspectos didácticos de la variable aleatoria han abordado, desde distintos enfoques teóricos, la manera que se debe enseñar este concepto a nivel escolar y universitario. Bizet y Ramos-Rodríguez (2019) indagan su comprensión en secundaria fundamentada en la teoría de Situaciones Didácticas (TSD) (Brousseau, 2007). Para ello aplican a 22 estudiantes chilenos de grado 10 (15 a 16 años) una propuesta de enseñanza contextualizada en un juego de azar, la cual requiere identificar y representar la variable aleatoria discreta y su función de probabilidad involucrada. Los resultados evidencian que 14 estudiantes reconocen y representan en diferentes registros la variable aleatoria: nueve en lenguaje verbal y gráfico (diagrama sagital), y cinco solo en lenguaje tabular. También 13 participantes identifican y representan la función de probabilidad: tres usan el lenguaje verbal, cinco utilizan tanto el lenguaje verbal como gráfico y los otros cinco emplean el lenguaje tabular. Las autoras entre sus conclusiones señalan que la propuesta diseñada favorece la comprensión del carácter funcional de la variable aleatoria dado que admite diversos registros de representación como gráfico, tabular y verbal.

En el contexto universitario, Kachapova (2012) y Kachapova y Kachapov (2012) proponen un enfoque para la enseñanza de la variable aleatoria. Entre sus resultandos presentan ideas para evitar o subsanar algunos errores típicos en torno al concepto, como por ejemplo: un suceso aleatorio es cualquier subconjunto del espacio muestral (verdadero solo para muestras finitas), una variable aleatoria es una función en el espacio muestral de valores reales, y esta es solo de tipo discreta o continua, obviando las mixtas. Además, exponen un enfoque holístico para la enseñanza de la variable aleatoria, definiéndola a través de la integral Riemann-Stieltjes y centrándose en una teoría general, no en detalles para las de tipo discretas y continuas. Los autores luego de aplicar varios años su propuesta, concluyen que el enfoque diseñado favorece la adquisición de conocimiento abstracto en torno al concepto y la diferenciación de las características específicas o generales de sus diferentes tipos.

También Maurer et al. (2020) diseñan e implementan dos actividades para la enseñanza de combinaciones lineales de variables aleatorias, una basada en un juego de azar y la otra en la resolución de problemas. La implementación se llevó a cabo en un curso de estadística con universitarios de Estados Unidos, quienes respondieron cuestionarios de evaluación antes y después de cada actividad y una encuesta de opinión. Los resultados sobre el aprendizaje de la variable aleatoria evidencian mejores puntuaciones finalizada la implementación de las dos actividades, aunque un leve mejor avance al concluir la resolución de los problemas. Además, en la encuesta todos los estudiantes opinaron que disfrutaron de las dos actividades. Finalmente, las autoras concluyen que el menor logro de aprendizaje observado al terminar la actividad del juego, puede deberse a que se invirtió bastante tiempo en la comprensión de sus normas, por ello proponen un video en que se expliquen estas.

Investigaciones sobre la didáctica de las distribuciones de probabilidad

Existen algunos trabajos centrados en elementos didácticos de las distribuciones de probabilidad, como el desarrollado por Kazak y Pratt (2017), quienes indagan en futuros profesores de matemáticas el uso de modelos probabilísticos para realizar inferencia estadística informal. Participaron 12 estudiantes turcos, los cuales debían crear una estrategia para ganar un juego empleando diferentes recursos (material concreto y software) y responder una entrevista. Los resultados muestran que el uso de material concreto dificulta identificar el espacio muestral y definir la variable aleatoria discreta involucrada, aunque luego de emplearse el software para simular el juego, se subsanan los errores y son reconocidos los valores de la variable aleatoria que poseen mayor posibilidad de ocurrencia. Así los estudiantes logran crear una estrategia adecuada para ganar el juego, dado que relacionan la distribución de los datos empíricos (distribución empírica) con la distribución teórica (modelo de probabilidad). Entre las conclusiones se afirma que tanto el uso de material concreto como la herramienta de simulación, aportan a la comprensión del juego y la realización de inferencia.

También hay investigaciones específicas sobre la didáctica del modelo binomial. Por ejemplo, Alvarado y Retamal (2014), diseñan un taller para la introducción del tema a nivel escolar (grados 11 y 12) por medio del análisis a experimentos aleatorios. Consideran como referente teórico herramientas del EOS y las diferentes representaciones de la binomial (manipulativo, computacional y algebraico). Entre los resultados obtienen 15 actividades que componen el taller, en las cuales se tienen en cuenta como conocimientos previos los conceptos de experimento aleatorio, simulación, variable aleatoria y función de probabilidad. A modo de conclusión los autores señalan que, el reto para la enseñanza del modelo binomial está en el diseño de actividades centradas en sus diferentes representaciones que permitan progresivamente introducir su significado dependiendo del grado de educación.

A diferencia de Bill et al. (2009) que elaboran e implementan una actividad para la enseñanza escolar del modelo binomial, la cual requiere cálculo de probabilidades y admite tres diferentes métodos de resolución: uso de la expresión de su función de probabilidad; aplicación del triángulo de Pascal; y realizar simulación con un software de datos. Participaron 19 estudiantes de 10 grado (15 a 16 años) quienes respondieron a dos cuestionarios de evaluación, antes y después de la actividad, y sus respuestas fueron analizadas desde la taxonomía SOLO (Structure of Observed Learning Outcome) propuesta por Biggs y Collis (1982). Entre los resultados obtenidos en el primer cuestionario se apreció dificultad para aplicar la regla del producto, mientras que, en el segundo, el 63% de los participantes logró aplicar los tres métodos estudiados. Así se concluye que el uso del software es más apreciado por los estudiantes menos habilidosos para las matemáticas, y las tareas intuitivas acerca de combinación de monedas son el origen para introducirse en la comprensión de la binomial.

Asimismo, García-García et al. (2018) y García y Hernández (2018) diseñan e implementan una actividad para la enseñanza escolar de la distribución binomial, por medio de un software de educación estadística. Participaron 46 estudiantes mexicanos de grado 11 y 12 (17 a 18 años), quienes respondieron un mismo cuestionario en dos instancias, previa y posterior aplicación de la actividad. Las respuestas fueron codificadas considerando los niveles de razonamiento de la taxonomía SOLO. Los resultados muestran un mejor desempeño de los estudiantes en la segunda instancia de aplicación del cuestionario, avance que se atribuye al desarrollo de la actividad basada en el uso de un simulador computacional. De esta manera se concluye que la utilización de tecnología favorece un aprendizaje relacionado con el contexto de los estudiantes.

Respecto a la distribución normal, es posible encontrar limitados estudios enfocados en sus aspectos didácticos. Por ejemplo, Prado y Gravoso (2011) crean, implementan y evalúan una unidad de aprendizaje para la enseñanza escolar de modelos de probabilidad (binomial, Poisson y normal), sustentada en la instrucción anclada (Crews et al., 1997). Participaron 64 estudiantes de Filipinas (13 a 15 años), quienes al comienzo y final de la unidad debieron responder un cuestionario y una encuesta para comparar sus aprendizajes y opinión, con otro grupo de estudiantes que aprendieron con una metodología tradicional. Los resultados evidencian que los estudiantes que aprendieron con base en la instrucción anclada obtuvieron significativamente mejores puntuaciones en los tres tópicos evaluados, en comparación con el otro grupo. Además, la encuesta constata que la unidad es entretenida, está contextualizada en situaciones reales, y fomenta el trabajo colaborativo y autónomo. Así se concluye que la intervención favoreció la mejora de habilidades en torno al razonamiento estadístico.

Posteriormente Salinas et al. (2018) diseñan e implementan secuencias didácticas para la enseñanza escolar de la normal, fundamentados en la metodología Lesson Study (Stigler y Hiebert, 1999) y la Aproximación Documental de lo Didáctico (Gueudet y Trouche, 2009). Para ello elaboraron dos secuencias de actividades: la primera usa un software y participaron 20 estudiantes (17 a 19 años); la segunda emplea hojas de cálculo y lápiz y se aplicó a 33 estudiantes (17 a 19 años). Finalizada cada implementación aplicaron un cuestionario de evaluación. Los resultados centrados en los profesores muestran que el conocimiento del contenido ha sido el más trabajado, y evidencia que la metodología usada fomenta el aprender a enseñar la normal y su mejoramiento. Además, la utilización de un software amplia el conocimiento de diferentes herramientas útiles para enseñar. A modo de conclusión proponen como mejora que el diseño de secuencias incluya guiones que contengan los objetivos de cada actividad, para así verificar su cumplimiento.

En el contexto universitario, C. K. Tan y C. P. Tan (2015) analizan la comprensión de la variable aleatoria y modelos de distribución (binomial, normal y Poisson), por medio de dos propuestas de enseñanza (una con calculadora y otra con papel y lápiz). Participaron 65 estudiantes (19 años) de Malasia distribuidos en dos grupos: grupo de experimentación (32), el cual trabajó con calculadora gráfica; grupo de control (33), quien aprendió de manera tradicional. También se aplicó un pre y post test de evaluación y consultó a los participantes su opinión respecto a la experiencia. Los resultados muestran que el grupo experimental logró significativamente mejores resultados en todos los tópicos, siendo mayor la disparidad en los contenidos de distribución binomial y normal. Respecto a su opinión, gran parte señaló que el uso de la calculadora facilitó la comprensión de los conceptos, pues se observa sus diferentes representaciones (gráfica y tabular) y disminuye la cantidad de cálculos, además promueve la participación activa y la confianza en sí mismo. Los autores concluyen que la tecnología es una buena herramienta para mejorar la comprensión de la probabilidad en estudiantes de bajo o alto rendimiento.

Por su parte, Tauber et al. (2004) diseñan e implementan una secuencia para la enseñanza universitaria de la distribución normal desde una perspectiva intuitiva, basada en el EOS. Participaron 143 estudiantes universitarios españoles que cursaban una asignatura de estadística y probabilidad, los cuales disponían de un software, papel y lápiz. Las respuestas a las actividades evidencian dificultades en: interpretar la probabilidad como el área bajo la curva normal; establecer los porcentajes asociados a la propiedad $\pm 3\sigma$; distinguir entre la distribución empírica y teórica; interpretar medidas de posición; diferenciar los estadísticos de los parámetros; e interpretar los valores de los coeficientes de curtosis y de asimetría. Entre las conclusiones se señala que para la enseñanza efectiva del modelo normal se requiere como conocimientos previos: probabilidad, distribución estadística, variable aleatoria y estadística, medidas de dispersión y centralización, entre otros.

También Hugues (2005) elabora e implementa una propuesta para enseñar la distribución normal a nivel universitario. El diseño de la propuesta se fundamenta en la investigación de Tauber (2001), así queda constituida por una serie de actividades en las que se utilizan recursos como hoja de papel y lápiz y una hoja de cálculo electrónica. Esta fue aplicada en 25 estudiantes mexicanos de ingeniería y sus resultados muestran que ellos logran reconocer: el vínculo entre los parámetros de la distribución normal y la forma de su curva; las condiciones que deben cumplir los parámetros de la distribución binomial para aproximarla a una normal; y las propiedades de la normal como su simetría. A modo de conclusión se señala que los participantes manifiestan dificultad en trabajar la corrección por continuidad.

Años más tarde Alvarado y Retamal (2010) diseñan e implementan una secuencia para enseñar en la universidad la aproximación de la binomial por la normal (Teorema de Laplace De-Moivre) fundamentados en el EOS. Participaron 104 estudiantes chilenos de una carrera de ingeniería, quienes finalizada la implementación respondieron un cuestionario de evaluación, donde sus respuestas muestran: sobre medir elementos que afectan a una buena aproximación del modelo binomial, un 16% se equivoca al suponer una mayor precisión de esta en los valores extremos del modelo; respecto a evaluar sensibilidad de los valores de los parámetros, un 18% responde de manera inadecuada, dado que confunden las condiciones de los parámetros n y p; y en torno a medir la aproximación de la binomial a la normal, el 60% contesta mal debido a la falta de distinción del valor de n para calcular la probabilidad, observándose el sesgo de insensibilidad al tamaño de la muestra. Entre las conclusiones se afirma que existen dificultades en realizar el procedimiento de estandarización y la corrección por continuidad.

2.2.3.2 Aprendizaje de la variable aleatoria y su distribución de probabilidad

En este apartado se presentan investigaciones que abordan el estudio de la variable aleatoria y sus distribuciones de probabilidad desde un ámbito cognitivo. Su característica principal es analizar estrategias desarrolladas por estudiantes al afrontar situaciones-problemas, con el propósito de identificar sus dificultades y/o errores, niveles de razonamiento, etc. Estos estudios se han llevado a cabo tanto en los niveles educativos escolares como en el contexto universitario.

Aprendizaje de la variable aleatoria en la educación escolar

Dentro del contexto escolar se han encontrado investigaciones centradas en observar el aprendizaje de la variable aleatoria en estudiantes, como García y Sánchez (2013), quienes indagan el desarrollo del razonamiento probabilístico en torno a esta (tipo discreta) y su distribución de probabilidad, sustentados en la taxonomía SOLO. Participaron 24 estudiantes mexicanos de grado 10 (15 a 16 años), los cuales debieron responder un cuestionario sobre un problema binomial. Los resultados se centran en las representaciones gráfica del modelo (gráfico de frecuencias relativas y gráfico de porcentajes) y evidencian que: el 75% de los participantes en ambos gráficos identifica el suceso aleatorio con mayor probabilidad; el 33% representa bien la distribución en los dos gráficos, manteniendo en ellos una proporción de sus resultados. Entre las conclusiones se afirma que los estudiantes logran reconocer que el valor central de la binomial posee mayor probabilidad que los extremos, aunque ellos manifiestan dificultad en representar consistencia entre los dos gráficos.

Luego, García et al. (2014) identifican componentes que caracterizan los niveles de razonamiento que manifiestan estudiantes escolares al resolver una tarea sobre la noción de variable aleatoria (discreta) y su distribución de probabilidad, basándose en la taxonomía SOLO. Las respuestas de 34 estudiantes de secundaria (grado 9, 14 a 15 años) y 46 de bachillerato (grado 10, 15 a 16 años) fueron analizadas y evidencian que globalmente tuvieron mejor desempeño los de mayor grado. Por ejemplo, en relación con la distribución de probabilidad, gran parte de los participantes reconoce que los valores de la variable aleatoria no poseen igual probabilidad, aunque el 59% de secundaria y el 52% de bachillerato lo fundamentan en factores no probabilísticos. Así se concluye que el 41% de repuestas provenientes de secundaria y el 12% de bachillerato están influencias por el sesgo de equiprobabilidad, es decir, suponer que los sucesos asociados a cualquier experimento aleatorio o fenómeno incierto, poseen igual probabilidad (Lecoutre, 1985).

Paralelamente, Flores et al. (2014) indagan en secundaria el desarrollo del razonamiento probabilístico en torno a la variable aleatoria y su distribución de probabilidad, fundamentados en la taxonomía SOLO. Para ello aplican a 69 estudiantes mexicanos de grado 8 (13 a 14 años) una actividad con apoyo de tecnología, además de un pre y post test de evaluación. Entre los resultados se observa sobre la relación de la variable aleatoria con el espacio muestral, que gran parte de los participantes, 82% en el pretest y 54% en el post test, manifiestan el sesgo de resultados aislados cuando identificaban algunos valores de la variable o elementos del espacio muestral, obviando que para analizar la probabilidad se considera una familia de experimentos aleatorios no únicamente un experimento. A modo de conclusión se afirma que el uso del software ayuda a disminuir razonamientos erróneos, como el sesgo de resultados aislados, relativo a suponer que cada repetición de un experimento no está relacionada con la anterior o posterior repetición (Konold, 1989).

Por su parte Salazar (2014) analiza en secundaria la construcción del concepto de variable aleatoria y su distribución de probabilidad con base en herramientas de la teoría APOE (Acción, Proceso, Objeto y Esquema) de Dubinsky (1991). Para ello lleva a cabo un estudio de tres casos, cuyo primer caso está integrado por cinco estudiantes chilenos de grado 9 a 12 (14 a 18 años), quienes debieron responder a un cuestionario. Los resultados revelan: reconocimiento parcial de la propiedad de la unión; falta de comprensión de axiomas fundamentales (la probabilidad de un suceso toma un valor entre 0 y 1); falta de reconocimiento de la variable aleatoria como una función; confusión entre los conceptos de suceso y variable aleatoria; y falta de identificación de la relación entre los elementos del espacio muestral y números reales. De esta forma, el autor concluye que los participantes no poseen una comprensión profunda de la variable aleatoria y su distribución de probabilidad.

Mientras que Guerrero et al. (2016) indagan en bachillerato el conocimiento sobre la esperanza de una variable aleatoria discreta a partir de la resolución de un problema. Participaron 63 estudiantes españoles: 30 de grado 11 (16 a 17 años) y 33 de grado 12 (17 a 18 años). En general los resultados muestran que los estudiantes del curso superior obtuvieron un mejor desempeño en las preguntas sobre identificar el valor esperado de una variable aleatoria, por ejemplo, en una de las interrogantes, solo el 30% de los participantes del menor curso respondió adecuadamente, a diferencia del curso superior que logró un 57,6% de respuestas buenas. Entre las conclusiones los autores señalan que los argumentos incorrectos observados se vinculan principalmente al sesgo de equiprobabilidad y cálculo de probabilidad.

Aprendizaje de las distribuciones binomial y normal en la educación escolar

En el contexto escolar también existen estudios desde una perspectiva cognitiva acerca de los modelos binomial y normal. Respecto al primero de ellos, Landín y Sánchez (2010) y Sánchez y Landín (2011) proponen niveles de razonamiento probabilísticos para evaluar en bachillerato la resolución de problemas binomiales. Sus trabajos se fundamentan en los niveles de desempeño del estudiante cuando responde una tarea, propuestos en la taxonomía SOLO, y en las componentes asociadas al conocimiento de la distribución binomial. Participaron 66 estudiantes mexicanos de grado 12 (17 a 18 años), quienes recibieron una instrucción previa sobre tópicos de probabilidad y posteriormente debieron resolver un cuestionario. Como resultado se obtiene una jerarquía de razonamiento en torno al modelo binomial, compuesta por cinco niveles crecientes de complejidad (1. subjetivo, 2. transicional, 3. cuantitativo-informal, 4. numérico y 5. abstracto), con posibilidad de adecuarse según los requerimientos específicos de una clase o necesidades de un investigador.

Posteriormente Landín (2013) y Sánchez y Landín (2014) describen el proceso en que escolares conocen y aplican la distribución binomial en la resolución de problemas, empleando como sustento teórico la jerarquía para el razonamiento binomial (Sánchez y Landín, 2011). Participaron 26 estudiantes mexicanos de grado 12 (17 a 18 años), los cuales respondieron a dos cuestionarios (uno previo y el otro posterior a un proceso de instrucción). A modo de conclusión se establece que, en el aprendizaje del modelo binomial, inicialmente el diagrama de árbol facilita la comprensión del concepto de combinatoria y la regla del producto, aunque para avanzar en la mejora de su aprendizaje se requiere la aplicación de estos últimos dos elementos. También se identificó que los problemas que más dificultó resolver fueron, en menor medida los relativos a combinatoria y comparación de probabilidad, mientras que los más complejos fueron los vinculados a la regla del producto y cálculo de probabilidad.

Del mismo modo, Sánchez y Carrasco (2018) y Sánchez et al. (2018) describen el razonamiento que manifiestan los escolares en la resolución de problemas que requieren la construcción de la distribución binomial, centrándose en sus argumentos en torno a la variable aleatoria, distribución, combinatoria y variabilidad. Participaron 34 estudiantes mexicanos de grado 12 (17 a 18 años) quienes respondieron dos cuestionarios. Los resultados evidencian dificultades relativas a: cálculo de probabilidades aplicando la función de distribución acumulada; representar el espacio muestral de un experimento aleatorio por medio de un diagrama de árbol; y determinar la probabilidad asociada a cada valor de la variable aleatoria, pues se apreció el sesgo de equiprobabilidad. Además, se reconoció que la comprensión de variable aleatoria, combinatoria y regla de producto son necesarias para lograr construir la distribución binomial.

A diferencia de Van Dooren et al. (2003) que indagan la presencia del sesgo de linealidad en el razonamiento de escolares al enfrentarse a situaciones que involucran la distribución binomial. Para tal efecto, elaboraron un cuestionario que aplicaron a 225 estudiantes belgas: 107 de grado 10 (15 a 16 años) y 118 de grado 12 (17 a 18 años). Los resultados evidencian que tanto los de mayor grado (87,5%) como los de menor edad (75,3%), suponen la existencia de una relación lineal entre la probabilidad de un suceso y los parámetros n, p o q. Esto se debe a que los participantes seleccionaron como respuestas alternativas que muestran una disminución o aumento de la probabilidad cuando disminuían o aumentaban dos parámetros involucrados, y manifiestan la relación a través de palabras, esquema o expresión algebraica, por ejemplo: $P = \frac{n \cdot p}{q}$ o variación de esta. Así se concluye la presencia del sesgo de linealidad (De Bock et al., 2002) que consiste en asignar propiedades de la función lineal a una situación u objeto matemático que no es posible.

Sobre el aprendizaje del modelo normal en la escuela, son limitados los estudios existentes. En este contexto González y Ojeda (2017) analizan la resolución de un problema luego de un proceso de instrucción sobre el tema, para lo cual utilizan las ideas estocásticas fundamentales de Heitele (1975) y la herramienta del triángulo epistemológico de Steinbring (2005). Participaron 26 estudiantes mexicanos de grado 12 (17 a 19 años), de los cuales el 31% respondió adecuadamente, el 54% incorrectamente y el 15% no presentó respuesta. Posteriormente se realizó una entrevista a dos de esos participantes, uno que presentó respuesta correcta al problema y otro que evidenció errores. En ella se constata que los estudiantes logran realizar el proceso de estandarización, pero no vincularlo con la curva normal y explicar su significado. De esta manera se concluye que existe una falta de reconocimiento del área bajo la curva que ha sido estandarizada y comprensión del proceso.

A continuación, Torres y Ojeda (2018), indagan la comprensión de los conceptos requeridos para el estudio de la función de densidad de la normal, basados en las ideas fundamentales de Heitele (1975) y el triángulo epistemológico de Steinbring (2005). Inicialmente cuatro estudiantes mexicanos de grado 12 (17 a 18 años) participaron en la implementación de una secuencia de enseñanza en torno al modelo normal y posterior resolución de un cuestionario de evaluación. Luego dos de ellos, con mejor desempeño, fueron seleccionados para una entrevista, observándose como resultado falencias en: distinguir entre continuidad de la función y cambio de su concavidad, en un intervalo, y diferenciar entre media aritmética y el valor de la variable aleatoria con mayor número de repeticiones en un conjunto de mediciones. Con respecto a esta última idea, los autores concluyen que es posible que genere dificultad para comprender el rol de la desviación estándar y la media en dicha función de densidad.

Por su parte, Valdez y Salinas (2019) exploran en bachillerato la resolución de problemas en torno a la normal fundamentados en la taxonomía SOLO. Así aplican un cuestionario a 53 estudiantes de grado 12 (17 a 18 años), organizados en dos grupos: 20 en el grupo A, los cuales aprendieron mediante un software; 33 en el grupo B, quienes recibieron instrucción con un enfoque tradicional. Entre los resultados obtienen una jerarquía de razonamiento en torno a la normal, específicamente para los tópicos de estandarización y cálculo de probabilidad, cada una constituida por cuatro niveles creciente de razonamiento (1. pre-estructural; 2. uni-estructural; 3. multi-estructural; y 4. relacional). Además, las respuestas al cuestionario manifiestan que en general el grupo B tuvo mejor desempeño que el grupo A, dado que la mayoría de sus respuestas se clasifican en el nivel de razonamiento más alto (relacional). A modo de conclusión se afirma que existe dificultad en relacionar el área bajo la curva normal con la probabilidad y se observa una correcta realización del procedimiento de estandarización, aunque existe falta de comprensión en algunas de sus etapas

Aprendizaje de la variable aleatoria en otros contextos

Dentro del ámbito universitario se han encontrado importantes estudios sobre los procesos cognitivos en torno al aprendizaje de la variable aleatoria. Algunos de ellos han sido desarrollados por Ruiz y Albert (2005) y Ruiz et al. (2006), quienes se interesan en explorar el proceso de construcción, concepciones y dificultades en torno a la variable aleatoria y su distribución de probabilidad, basándose en TSD (Brousseau, 1997). Participaron dos estudiantes mexicanos de una carrera de ciencias sociales, los cuales respondieron en pareja un problema y luego fueron entrevistados. Los resultados evidencian dificultades en: identificar la aleatoriedad en el problema, pues se intenta crea una solución determinista; reconocer que las probabilidades calculadas corresponden a valores de la función de probabilidad; identificar a la variable aleatoria como una función; representar gráficamente su distribución de probabilidad; y trabajar con una función discontinua, pues se asocia a ella una gráfica cuyo dibujo es continuo.

Después, Ruiz et al. (2011) evalúan la comprensión de la variable aleatoria y su vinculación con la variable estadística, a través de la realización de un proyecto abierto, donde participaron 101 futuros profesores españoles de educación primaria. Los resultados evidencian que solo un 33% de ellos logró analizar la distribución de la variable estadística y realizar inferencias y obtener conclusiones parciales en relación con la variable aleatoria, mientras que solo un 4% de los estudiantes pudo finalizar completamente el proceso de inferencia estadística, constatándose que es una tarea que les resulta difícil. A modo de conclusión los autores señalan que existen inconvenientes en el desarrollo de tareas sobre modelación, lo cual en este estudio afectó en lograr vincular la variable estadística con la aleatoria. Además, se observó dificultad en el razonamiento en torno a la variabilidad de la variable aleatoria.

A diferencia de Pérez y Parraguez (2013) que analizan la comprensión y diferenciación de los conceptos de aleatoriedad y determinismo para la construcción de la variable aleatoria, usando la teoría APOE. Por ello elaboran una descomposición genética hipotética del objeto (esquema que describe la comprensión y vínculo de diversos conceptos que requiere lograr un sujeto para construir un objeto matemático). Con el propósito de validar esta descomposición, diseñan un cuestionario aplicado a nueve futuros profesores chilenos de matemáticas. Entre sus principales resultados afirman que la diferenciación entre los conceptos suceso y variable es fundamental, es decir, identificar qué se está investigando (suceso) y a través de qué característica se hará (variable), pues si el estudiante no comprende estos temas no logrará modelar una situación-problema por medio de una variable aleatoria.

Al mismo tiempo, Fernández et al. (2013) diseñan e implementan una secuencia de enseñanza para la construcción del significado de variable aleatoria, fundamentados en la alfabetización, razonamiento y pensamiento estadístico (Ben-Zvi y Garfield, 2004). Participó un grupo de futuros profesores colombianos de matemáticas, quienes respondieron a dos talleres, el primero sobre espacio muestral y el segundo referente a la variable aleatoria. El análisis a las respuestas muestra dificultad en identificar el experimento aleatorio y expresar el espacio muestral de manera formal (utilizando notación de conjunto). Además, se evidencian errores como visualizar los valores de la variable aleatoria como resultado de un experimento aleatorio.

Luego, Calandra y Costa (2015) analizan en universitarios dificultades en torno a la variable aleatoria (continua) y su distribución de probabilidad, fundamentadas en la teoría Antropológica de lo Didáctico (Chevallard, 1999). Participaron 50 estudiantes argentinos de ingeniería que cursaban una asignatura sobre tópicos estocásticos. Inicialmente se analiza la manera en que se introducen los conceptos en las guías de contenido y libros de texto del curso, cuyos resultados muestran discrepancias entre la presentación del contenido y los ejercicios propuestos. Por ejemplo, algunos libros destacan el vínculo de los histogramas con la función de densidad, pero no se presentan problemas en que se refleje este. Luego se aplica un cuestionario y sus resultados evidencian dificultades en: identificar la variable aleatoria en el problema; resolución mecánica y descontextualizada; y falta de interpretación del significado de la función de densidad en el contexto del problema.

Aprendizaje de las distribuciones binomial y normal en otros contextos

También a nivel universitario existen estudios acerca de la comprensión de modelos probabilísticos. En particular sobre el modelo binomial, Vilca (2015) identifica errores al abordar un problema vinculado a este, basándose en el EOS. Para ello aplicó un cuestionario a 18 estudiantes peruanos (16 a 19 años) de una carrera de administración. Los resultados muestran que: ninguno logró definir la variable aleatoria involucrada en el problema; sólo el 17% determinó el recorrido de la variable aleatoria discreta y el 33% presentó errores en el lenguaje matemático al expresar este; y el 22% identificó incorrectamente los valores de los parámetros de la distribución binomial (n y p) y el 5% se equivocó en la simbología utilizan para señalar estos. Así, se concluye dificultad en reconocer tanto la variable aleatoria como los parámetros del modelo binomial.

Por su parte, Galicia et al. (2013) indagan el razonamiento al resolver problemas vinculados al modelo binomial, basándose en la taxonomía SOLO. Participaron 18 profesores en formación y 59 estudiantes mexicanos de grado 11 y 12 (16 a 18 años), los cuales respondieron a un cuestionario En general los resultados evidencian que los futuros profesores presentan un mejor desempeño en comparación a los escolares. Los autores concluyen que tanto el diagrama de árbol como el enfoque de probabilidad clásica favorece la identificación del modelo binomial en problemas de distribución, pero cuando el parámetro (número de ensayos) toma un valor grande, es decir $n \geq 8$, se requiere la utilización de la regla del producto en vez del diagrama de árbol.

Mientras que Alvarado et al. (2018) analizan intuiciones y heurísticas relacionadas a la probabilidad, fundamentados en sus diferentes enfoques e ideas acerca del razonamiento probabilístico de Tversky y Kahneman (1980). Participaron 257 estudiantes chilenos de una carrera de ingeniería que cursaban una asignatura de estadística, los cuales respondieron a dos cuestionarios, uno al comienzo y otro al final de esta. Los resultados muestran sobre la distribución binomial que inicialmente el 7% de ellos responde bien, pero luego del proceso de instrucción se constató una mejora de su desempeño (68,2%). Además, se observó dificultades en definir la variable aleatoria, identificar el modelo binomial y sus parámetros. Así mismo se presentó el sesgo de insensibilidad al tamaño de la muestra, que consiste en suponer que una muestra de tamaño pequeño es igual de representativa en las características estadísticas que la población donde surge (Tversky y Kahneman, 1971).

Luego, Toledo et al. (2019) analizan el razonamiento de profesores al abordar problemas de probabilidad de tipo binomial, con base en la taxonomía SOLO. Participaron 63 profesores chilenos de matemáticas quienes respondieron a un cuestionario que incluyó situaciones equiprobables y no equiprobables. Los resultados muestran sobre la situación equiprobable $\left(\mathrm{p} = \frac{1}{2}\right)$ que, el 68,25% de ellos responde bien, empleando el diagrama de árbol y la probabilidad clásica, y solo el 7,98% lo hace aplicando la binomial. En la situación no equiprobable $\left(\mathrm{p} = \frac{5}{8}\right)$, el 32,75% responde correctamente usando combinatoria y regla de Laplace o técnicas de conteo y solo el 14,29% lo realiza utilizando la binomial. Entre las conclusiones se destaca la utilización de la fórmula binomial, aunque no se aprecia el uso de un lenguaje formal, identificación de sus parámetros y definición de la variable aleatoria involucrada.

En relación con la comprensión de la distribución normal en la universidad, autores como Batanero et al. (2001) abordan esta temática, basándose en el EOS. Para ello aplican a 55 estudiantes universitarios españoles de diferentes carreras (educación primaria, psicopedagogía, psicología o economía) un cuestionario y una prueba. Los resultados en el cuestionario muestran que alrededor del 70% respondió adecuadamente, aunque se observó que les resulta difícil: asociar a la normal una expresión algebraica; notar que el modelo normal toma valores negativos y positivos; interpretar el área de un histograma; reconocer que el área total bajo una curva normal es igual a una unidad; calcular valores tipificados e inversos; y distinguir entre la distribución empírica y teórica. En las respuestas a la prueba se observaron dificultades en torno a la normal, posteriormente constatadas por Batanero et al. (2004), estas son: interpretar los coeficientes de curtosis y de asimetría; identificar las condiciones que debe cumplir una distribución de variable aleatoria discreta para aproximarla a una normal; y comprender los conceptos de parámetros y estadísticos.

Posteriormente Alvarado y Batanero (2007) indagan la comprensión de la aproximación binomial a la normal fundamentados en el EOS. Así diseñan e implementan una secuencia de enseñanza y aplican un cuestionario de evaluación. Participaron 123 universitarios chilenos estudiantes de ingeniería y sus respuestas evidencian errores como: confusión en la condición de los parámetros n y p para la aproximación (16%); afirmar que en valores pequeños de n se obtiene una buena aproximación (5%); falta de reconocimiento de la relevancia de los valores de los parámetros de la binomial para la rapidez de su convergencia a la normal (31%); suponer una mayor precisión en la aproximación a la normal en los valores extremos de la binomial (24,7%); equivocación en los límites de la corrección por continuidad (21%); y confusión entre la expresión para calcular la media y varianza. Entre las conclusiones se observa la presencia del sesgo de insensibilidad al tamaño de la muestra (55%).

A continuación, Carpio et al. (2009) reconocen elementos que intervienen en la enseñanza y aprendizaje de la normal con base en herramientas del EOS. Para ello aplicaron a 45 estudiantes peruanos de una carrera de administración y economía un cuestionario de evaluación. Las respuestas de los participantes muestran que les dificulta aplicar la propiedad de reproductividad (si se suman variables aleatorias con distribución normal se obtendrá como resultado una variable distribuida normalmente), y confunden la varianza con la desviación estándar en el proceso de estandarización. De esta manera se concluye que el 35% de los estudiantes se equivoca al definir una nueva variable aleatoria usando notación de sumatoria.

A diferencia de Bansilal (2012; 2014) que explora en profesores la comprensión de la normal por medio de la resolución de problemas, fundamentada en la teoría de Registro de Representaciones Semiótica (Duval, 2006) y en la teoría APOE. Participaron 290 profesores de matemáticas sudafricanos que imparten docencia en escuelas, quienes en el contexto de un curso de perfeccionamiento respondieron a un cuestionario. Los resultados de las investigaciones muestran respuestas correctas sobre: realizar un proceso de estandarización (27%); llevar a cabo dos procesos de estandarización (14%); y determinar un valor desconocido a partir de una probabilidad dada (16%). Así se concluye que los profesores poseen un conocimiento limitado de la normal y se propone trabajar con un software que podría favorecer la observación de sus propiedades en diferentes registros de representación.

Además, González et al. (2018) analizan la comprensión de profesores sobre la normal, con base en las ideas estocásticas de Heitele (1975) y el triángulo epistemológico de Steinbring (2005). Participaron dos profesores de matemáticas mexicanos con experiencia en docencia universitaria, los cuales respondieron paralelamente un problema acerca de la normal y una entrevista. En los resultados se observa que uno de ellos presenta inconvenientes en identificar el espacio muestral y los valores de la función de densidad asociada al problema, mientras que, al otro le dificulta graficar la normal estándar. Además, ambos no logran reflexionar sobre el proceso de estandarización. A modo de conclusión se menciona que los participantes poseen problemas para diferenciar entre el significado frecuencial y clásico de la probabilidad, vincular la expresión $\mu \pm \sigma$ con puntos de inflexión de la curva normal y reflexionar en torno a los cálculos llevados a cabo.

2.2.4 Conclusiones e implicaciones didácticas

En este trabajo se analizaron las principales investigaciones que reporta la literatura en el campo de educación matemática sobre la enseñanza y aprendizaje de la variable aleatoria y su distribución de probabilidad, tanto a nivel escolar como universitario. Por ello se constata que, se ha encontrado una menor cantidad de estudios que abordan sus aspectos epistemológicos y didácticos (enseñanza) en comparación con los que tratan sus elementos cognitivos (aprendizaje).

Con relación al ámbito epistemológico, se aprecian investigaciones que han identificado de manera superficial elementos de la variable aleatoria (problemas, procedimientos, diferentes representaciones, etc.) necesarios para su enseñanza en la escuela (Ortiz, 2002) y universidad (Ruiz, 2013). Respecto al modelo binomial, dentro de lo que se ha buscado, no se han hallado estudios que entreguen directrices sobre los elementos principales a considerar en su enseñanza. Contrario a lo ocurrido con la normal, los resultados obtenidos son más alentadores, dado que se encontró detalladamente los elementos fundamentales a incluir en su proceso de enseñanza, tanto en la educación superior (Tauber, 2001) como escolar obligatoria (Valverde, 2017).

Sobre los aspectos didácticos se descubrieron similar cantidad de trabajos dirigidos a la educación universitaria y escolar. Estos principalmente se centran en propuestas para la enseñanza superior de la variable aleatoria (C. K. Tan y C. P. Tan, 2015; Kachapova, 2012; Kachapova y Kachapov, 2012; Kazak y Pratt, 2017; Maurer et al., 2020) y el modelo normal (Alvarado y Retamal, 2010; C. K. Tan y C. P. Tan, 2015; Hugues, 2005; Tauber et al., 2004). Además de la enseñanza a nivel escolar de la binomial (Alvarado y Retamal, 2014; Bill et al., 2009; García-García et al., 2018; García y Hernández, 2018).

Mientras que existe una menor cantidad de propuestas de instrucción en el contexto escolar en torno a la variable aleatoria (Bizet y Ramos-Rodríguez, 2019) y la normal (Prado y Gravoso, 2011; Salinas et al., 2018), lo mismo que ocurre en el ámbito universitario con la binomial (C. K. Tan y C. P. Tan, 2015).

Entre los trabajos de diseño de actividades se evidenció que el recurso de la tecnología, por medio de una calculadora gráfica o software de simulación de datos, es el que ha permitido obtener un mejor rendimiento de los estudiantes (Bill et al., 2009; C. K. Tan y C. P. Tan, 2015; García-García et al., 2018; García y Hernández, 2018; Kazak y Pratt, 2017; Salinas et al., 2018; Tauber et al., 2004). En este contexto se cree importante elaborar actividades para abordar la variable aleatoria y su distribución de probabilidad a través de la modelización, que en términos de Garfield et al. (2008) hace referencia al uso de la simulación para generar datos, y estimación de probabilidades mediante dispositivos aleatorios y herramientas de simulación. Dado que podría favorecer la enseñanza de los temas analizados y permitiría vincular la variable estadística (distribución de datos) con la aleatoria (distribución de probabilidades), proceso importante para realizar inferencia estadística informal, es decir, *"forma en que el estudiante utiliza conocimiento estadístico con el propósito de elaborar argumentos para apoyar inferencias sobre poblaciones desconocidas, basados en nuestras observaciones*" (Zieffler et al., 2008, p. 44).

Dentro de los estudios desarrollados desde una perspectiva cognitiva, se hallaron alrededor de la misma cantidad de investigaciones que abordan la comprensión de la variable aleatoria y su distribución en la educación escolar y universitaria. En particular, se han encontrado importantes resultados sobre el razonamiento probabilístico en torno a estos conceptos. Por una parte, en el contexto escolar los niveles de comprensión del modelo binomial (Landín y Sánchez, 2010; Sánchez y Landín, 2011) y modelo normal (Valdez y Salinas, 2019). Sin embargo, en la literatura indagada existen vacíos que serían importantes investigar, por ejemplo, analizar de manera articulada el nivel de comprensión de la variable aleatoria y los modelos de probabilidad al término de la educación escolar, o indagar el nivel de comprensión de la distribución binomial en la educación superior.

Por otra parte, se ha reconocido la presencia de los siguientes sesgos cognitivos en el razonamiento de estudiantes: el sesgo de equiprobabilidad (Sánchez et al., 2018); el sesgo de resultados aislados (Flores et al., 2014); el sesgo de linealidad (Van Dooren et al., 2003); y el sesgo de insensibilidad al tamaño de la muestra (Alvarado et al., 2018).

Así mismo sobre los tópicos indagados se constataron a nivel escolar diversas dificultades como: identificar el dominio de la variable aleatoria (Bizet y Ramos-Rodríguez, 2019) y su recorrido (García et al., 2014), reconocer a la variable aleatoria como una función (Salazar, 2014); calcular el valor esperado de una variable aleatoria discreta (Guerrero et al., 2016); aplicar la función de probabilidad de la binomial (Sánchez y Landín, 2014) y su función de distribución (Sánchez y Carrasco, 2018); y comprender etapas del procedimiento de estandarización y relacionar el área bajo la curva normal con la probabilidad (Valdez y Salinas, 2019), etc.

Mientras que en los estudiantes universitarios ha sido posible verificar dificultades, tales como: identificar la variable aleatoria como función y representar gráficamente su distribución de probabilidad (Ruiz y Albert, 2005); vincular la variable estadística con la aleatoria (Ruiz et al., 2011); reconocer los parámetros de la distribución binomial (Toledo et al., 2019) y determinar su recorrido (Vilca, 2015); interpretar la probabilidad como el área bajo la curva normal y establecer los porcentajes asociados a la propiedad $\pm 3\sigma$ (Tauber et al., 2004); diferenciar entre la distribución empírica y teórica y comprender los conceptos de parámetros y estadísticos (Batanero et al., 2001); reconocer las condiciones de los parámetros n y p para emplear la aproximación de la binomial a la normal (Batanero et al., 2004) y aplicar la corrección por continuidad (Alvarado y Batanero, 2007), etc.

A partir de este estudio, se ha constatado la relevancia de la variable aleatoria y su distribución de probabilidad dentro de la educación estocástica. Por consiguiente, es fundamental su inclusión en todo curso básico de estadística y probabilidad. En su aprendizaje intervienen diversos conceptos básicos tales como: espacio muestral, función, suceso aleatorio, experimento aleatorio, etc., además de involucrar los enfoques clásico y frecuencial de la probabilidad. Todos ellos son necesarios de comprender previamente en la educación escolar, pero para los estudiantes tienen sus dificultades propias, que luego al integrarlos y relacionarlos requieren de un mayor esfuerzo cognitivo. El limitado conocimiento de dichos elementos genera posteriormente dificultades en la comprensión de la variable aleatoria, distribución binomial y/o distribución normal, así como también la presencia de sesgos cognitivos en torno a estos tópicos, los cuales se manifiestan desde finales de la educación escolar, durante la formación universitaria y persisten en profesores de matemáticas en ejercicio. Por tanto, es fundamental fortalecer la formación docente en torno a los temas en cuestión, incluyendo sus aspectos epistemológicos, didácticos y cognitivos, de esta manera proporcionarles las herramientas suficientes para enfrentar con éxito la labor de su enseñanza.

También es necesario que la enseñanza considere los resultados de este estudio con el propósito de insistir en la comprensión de la variable aleatoria y las distribuciones de probabilidad. Según se expuso en la introducción las directrices curriculares españolas (MECD, 2015) en el eje estadística y probabilidad incluyen estos tópicos, enfatizando en las distribuciones binomial y normal, y la aproximación de la binomial por la normal. Aunque en este contexto, en bachillerato podría incluirse abordar la distribución empírica por medio de softwares de simulación de datos para introducir las distribuciones teóricas y relacionar ambas, de esta manera mostrar de forma empírica como la frecuencia relativa convergen a la probabilidad teórica. Además, ello permitiría vincular la variable estadística con la aleatoria, por medio de los enfoques frecuencial y clásico de la probabilidad, pues como ha quedado de manifiesto la tecnología provee recursos didácticos que favorece la comprensión de los tópicos de interés.

Finalmente, la recopilación de estudios enfocados en la comprensión de la variable aleatoria y su distribución de probabilidad, se cree que podría aportar a mejorar el proceso de su enseñanza y aprendizaje, a la identificación de dificultades en los estudiantes y al diagnóstico de los posibles sesgos que ellos pueden manifestar. De esta forma, se espera que estos resultados contribuyan a fortalecer el proceso de diseño instruccional llevado a cabo por el profesorado, con el fin de ayudar a los estudiantes a subsanar dificultades potenciales en el aprendizaje de dichos temas.

Referencias

Alvarado, H. y Batanero, C. (2007). Dificultades de Comprensión de la Aproximación Normal a la Distribución Binomial. *Números, revista de didáctica de las matemáticas*, 67, 1-7.

Alvarado, H., Estrella, S., Retamal, L. y Galindo, M. (2018). Intuiciones probabilísticas en estudiantes de ingeniería: implicaciones para la enseñanza de la probabilidad. *RELIME*, *21*(2), 131-156.

Alvarado, H. y Retamal, L. (2010). La aproximación binomial por la normal: Una Experiencia de Reflexión Sobre la Práctica. *Revista Paradigma*, *31*(2), 89-108.

Alvarado, H. y Retamal, M. (2014). Representaciones de la distribución de probabilidad binomial. En L. Andrade (Ed.), *Memorias del I Encuentro Colombiano de Educación Estocástica* (pp. 98-109). Asociación Colombiana de Educación Estocástica.

Bansilal, S. (2012). Using conversions and treatments to understand students' engagement with problems based on the normal distribution curve. *Pythagoras, 33*(1), 1-13.

Bansilal, S. (2014). Using an apos framework to understand teachers' responses to questions on the normal distribution. *Statistics Education Research Journal*, *13*(2), 42-57.

Batanero, C., Tauber, L. y Sánchez, V. (2001). Significado y comprensión de la distribución normal en un curso introductorio de análisis de datos. *Quadrante, 10*(1), 59-92.

Batanero, C., Tauber, L. y Sánchez, V. (2004). Students' reasoning about the normal distribution. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 257–276). Springer.

Ben-Zvi, D. y Garfield, J. (2004). Statistical literacy, reasoning and thinking: Goals, definitions andchallenges. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3- 16). Springer.

Biggs, J. y Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. Academic Press.

Bill, A., Watson, K. y Gayton, P. (2009). Guessing Answers to Pass a 5- item True False Test: Solving a Binomial Problem Three Different Ways. En R. Hunter, B. Bicknell y T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conferencia del Grupo de Investigación en Educación en Matemáticas de Australasia* (pp. 57-64). MERGA.

Bizet, V. y Ramos, E. (2019). Una experiencia de enseñanza para abordar la variable aleatoria con estudiantes de secundaria. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística* (pp. 1-10). Departamento de didáctica de la matemática de la Universidad de Granada.

Brousseau, G. (1997). *Theory of didactical situations in mathematics. Didactique des mathématiques, 1970-1990*. Kluwer.

Brousseau, G. (2007). *Iniciación al estudio de la teoría de las situaciones didácticas*. Libros del Zorzal.

Calandra, M. y Costa, V. (2015). La problemática de la enseñanza y aprendizaje del concepto de variable aleatoria continua y de función de densidad de probabilidad. En FaHCE (Ed.), *Actas IV jornadas de enseñanza e investigación educativa en el campo de las ciencias exactas y naturales* (pp. 1–10). UNLP.

Carpio, M., Gaita, C., Wilhelmi, M. y Sáenz de Cabezón, A. (2009). Significados de la distribución normal en la universidad. En M. González, M. González y J. Murillo (Eds.), *Investigación en Educación Matemática XIII* (pp. 1-15). SEIEM.

Chevallard, Y. (1999). El análisis de las prácticas docentes en la teoría antropológica de los didáctico. *Recherches en Didáctique des Mathématiques, 19*(2), 221-266.

Choo-Kim, T. y Choo-Peng, T. (2015). Effects of the handheld technology instructional approach on performances of students of different achievement levels. *Computers and Education*, *82*, 306-314.

Crews, T., Biswas, G., Goldman, S. y Bransford, J. (1997). Anchored interactive learning environments. *International Journal of Artificial Intelligence in Education*, *8*, 142-178.

De Bock, D., Verschaffel, L. y Janssens, D. (2002). The effects of different problem presentations and formulations on the illusion of linearity in secondary school students. *Mathematical Thinking and Learning*, *4*(1), 65–89.

Dinges, H. (2005). Variables, in particular random variables. En M. Hoffmann, J. Lenhard, y F. Seeger (Eds.), *Activity and sign grounding mathematics education* (pp. 305–311). Springer.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. En D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 95-126). Kluwer.

Duval, R. (2006). A cognitive analysis of problems of comprehension in the learning of mathematics. *Educational Studies in Mathematics*, *61*, 103–131.

Fernández, F., Andrade, L. y Sarmiento, B. (2013). Rehaciendo el camino hacia la comprensión de la variable aleatoria. En P. Perry, C. Samper, Ó. Molina, L. Camargo, A. Echeverry, F. Fernández, B. Sarmiento (Eds.), *Aportes investigativos para el diseño curricular en geometría y estadística* (pp. 93–169). Universidad Pedagógica Nacional.

Flores, B., García, J. y Sánchez, E. (2014). Avances en la calidad de las respuestas a preguntas de probabilidad después de una actividad de aprendizaje con tecnología. En M. González, M. Codes, D. Arnau y T. Ortega (Eds.), *Investigación en Educación Matemática XVIII* (pp. 307-316). SEIEM.

Fox, D. (1981). *El proceso de investigación en la educación*. Eunsa.

Franklin, C., Kader, G., Newborn, D., Moreno, J., Peck, R., Perry, M. y Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: a pre-k–12 curriculum framework.* ASA.

Galicia, S., Nájera, A. y Sánchez, E. (2013). Niveles de razonamiento frente a problemas binomiales. En J. M. Contreras, G. Cañadas, M. Gea y P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria* (pp. 409-416). Departamento de didáctica de la matemática de la Universidad de Granada.

García-García, J., Arredondo, E. y Márquez, M. (2018). Desarrollo de la noción de distribución binomial en estudiantes de bachillerato con apoyo de tecnología. *Revista Paradigma*, *39*(2), 92-106.

García-García, J. y Hernández, E. (2018). Niveles de razonamiento probabilístico de estudiantes de bachillerato sobre la noción de la distribución binomial. En L. Serna y D. Páges (Eds.), *Acta Latinoamericana de Matemática Educativa* (pp. 962-969). CLAME.

García, J., Medina, M. y Sánchez, E. (2014). Niveles de razonamiento de estudiantes de secundaria y bachillerato en una situación-problema de probabilidad. *Avances de Investigación en Educación Matemática*, *6*, 5–23.

García, J. y Sánchez, E. (2013). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a una situación básica de variable aleatoria y distribución. En J. M. Contreras, G. R. Cañadas, M. M. Gea, y P. Arteaga (Eds.), *Actas de las jornadas virtuales de didáctica de la estadística, probabilidad y combinatoria* (pp. 417–424). Departamento de didáctica de la matemática de la Universidad de Granada.

Garfield, J., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C. y Zieffler, A. (2008). Learning to reason about statistical models and modeling. En J. Garfield y D. Ben-Zvi, *Developing students' statistical reasoning, connecting research and teaching practice* (pp. 143-163). Springer.

Godino, J., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM Mathematics Education, 39*, 127-135.

González, Y. y Ojeda, A. (2017). Comprensión de la distribución normal en bachillerato. En L. Serna (Ed.), *Acta Latinoamericana de Matemática Educativa, 30* (pp. 207-2017). CLAME.

González, J., Ojeda, A. y Palacios, J. (2018). Comprensión de profesores de la distribución normal. En D. García e I. Pérez (Eds.), Acta Latinoamericana de Matemática Educativa, 31(2), (pp. 1764-1772). Ciudad de México: CLAME.

Guerrero, H., Batanero, C. y J. M. Contreras. (2016). Conocimientos sobre esperanza matemática en alumnos de bachillerato. En F. España (Ed.), *Actas del XVI congreso de enseñanza y aprendizaje de las matemáticas* (pp. 26-35). Sociedad Andaluza de Educación Matemática THALES.

Gueudet, G. y Trouche, L. (2009). Towards new documentation systems for mathematics teachers? *Educational Studies in Mathematics*, *71*, 199–218.

Heitele, D. (1975). An epistemological view of fundamental stochastic ideas. *Educational Studies in Mathematics*, *6*, 187–205.

Hernández, R., Fernández, C. y Baptista, L. (2014). Metodología de la investigación (6.a ed.). McGrawHill.

Hugues, H. (2005). Uso de Hojas Electrónicas en la Enseñanza de la Distribución Norma. En J. Lezama, M. Sánchez y J. Molina (Eds.), *Acta Latinoamericana de Matemática Educativa, 18* (pp. 757–763). CLAME

Kachapova, F. (2012). A general approach to teaching random variables. *Mathematics teaching-research journal online*, *5*(2), 1-16.

Kachapova, F. y Kachapov, I. (2012). Students' misconceptions about random variables. *International Journal of Mathematical Education in Science and Technology*, *41*(3), 963-971.

Kazak, S. y Pratt, D. (2017). Pre-service mathematics teachers' use of probability models in making informal inferences about a chance game. *Statistics Education Research Journal*, *16*(2), 287-304.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, *6*, 59-98.

Landín, P. (2013). Niveles de razonamiento probabilístico de estudiantes de bachillerato sobre problemas binomiales En J. M. Contreras, G. Cañadas, M. Gea y P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística*, *Probabilidad y Combinatoria* (pp. 425-431). Departamento de didáctica de la matemática de la Universidad de Granada.

Landín, P. y Sánchez, E. (2010). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a tareas de distribución binomial. *Educación Matemática Pesquisa, 12*(3), 598-618.

Lecoutre, M. (1985). Effect d' informations de nature combinatoire et de nature fréquentielle sur le judgements probabilistes. *Recherches en Didactique des Mathématiques, 6*, 193- 213.

Maurer, K., Hudiburgh, L. y Werwinski, L. (2020). What do students gain from games? Dice gamesvs word problems. *Teaching Statistics, 42*(2), 41-46.

Ministerio de Educación, Cultura y Deporte. (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato.* Boletín Oficial del Estado.

National Council of Teachers of Mathematics. (2002). *Principles and standards for school mathematics*. NCTM.

Ortiz, J. (2002). *La probabilidad en los libros de texto*. Grupo de Investigación en Educación Estadística de la Universidad de Granada.

Pérez, B. y Parraguez, M. (2013). Construcciones mentales de los conceptos aleatorio y determinista partir de la regresión lineal. En R. Flores (Ed.), *Acta Latinoamericana de Matemática Educativa, 26* (pp. 589–598). CLAME.

Prado, M. y Gravoso, R. (2011). Improving high school students' statistical reasoning skills: a case of applying anchored instruction. *The Asia-Pacific Education Researcher, 20*(1), 61-72.

Ruiz, B. (2013). *Análisis epistemológico de la variable aleatoria y comprensión de objetos matemáticos relacionados por estudiantes universitarios* [Tesis de Doctorado, Universidad de Granada].

Ruiz, B. y Albert, J. (2005). Didáctica de la Probabilidad y Estadística El Caso de la Variable Aleatoria. En J. Lezama, M. Sánchez y J. Molina (Eds.), *Acta Latinoamericana de Matemática Educativa, 18* (pp. 185- 191). CLAME.

Ruiz, B., Albert, J. y Batanero, C. (2006). An exploratory study of students' difficulties with random variable. En A. Rossman y B. Chance (Eds.), *7 International Conference on Teaching Statistics* (pp. 1-6). IASE.

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*(39), 431–449.

Salazar, R. (2014). *La variable aleatoria con probabilidad desde la perspectiva de la teoría APOE* [Tesis de Máster, Pontificia Universidad Católica de Valparaíso].

Salinas, J., Valdez, J. y Salinas-Hernández, U. (2018). Un acercamiento a la metodología lesson study para la enseñanza de la distribución normal. En L. Rodríguez-Muñiz, L. Muñiz-Rodríguez, A. Aguilar-González, P. Alonso, F. García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 525- 534). SEIEM.

Sánchez, E. y Carrasco, G. (2018). El razonamiento probabilístico de estudiantes de bachillerato en actividades de distribución binomial. En L. Rodríguez-Múniz, L. Múñiz-Rodríguez, Á. Aguilar-González, P. Alonso, F. García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 535- 543). SEIEM.

Sánchez, E., Carrasco, G. y Herrera, M. (2018). Fundamental ideas in the probabilistic reasoning of high - school students in binomial distribution activities. En M. Sorto, A. White, y L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics* (pp.1-6). IASE.

Sánchez, E. y Landín, P. (2011). Fiabilidad de una jerarquía para evaluar el razonamiento probabilístico acerca de la distribución binomial. En M. Marín, G. Fernández, L. Blanco y M. Palarea (Eds.), *Investigación en Educación Matemática XV* (pp. 533- 541). SEIEM.

Sánchez, E. y Landín, P. (2014). Levels of probabilistic reasoning of high school students about binomial problems. En E. Chernoff y B. Sriraman (Eds.), *Probabilistic thinking presenting plural perspectives* (pp. 581-597). Springer.

Steinbring, H. (2005). *The construction of new mathematical knowledge in classroom interaction. An epistemological perspective*. Springer.

Stigler, J. y Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* The Free Press.

Tauber, L. (2001). *La construcción del significado de la distribución normal a partir de actividades de análisis de datos* [Tesis de Doctorado, Universidad de Sevilla].

Tauber, L., Sánchez, V. y Batanero, C. (2004). Diseño, implementación y análisis de una secuencia de enseñanza de la distribución normal en un curso universitario. *Revista EMA, 9*(2), 82-105.

Toledo, Á., Montenegro, D. y Vicencio, I. (2019). Niveles de razonamiento frente a problemas binomiales. *Brazilian Journal of Development, 5*(6), 5399- 5410.
Torres, O. y Ojeda, A. (2018). Requisitos conceptuales de la función de densidad normal como modelo de la realidad. En D. García e I. Pérez (Eds.), *Acta Latinoamericana de Matemática Educativa, 31(2),* (pp. 1085-1093). CLAME.
Tversky, A. y Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*(2), 105–110.
Tversky, A. y Kahneman, D. (1980). Causal schemas in judgments under uncertainty. En E. Fishbein (Ed.), *Progress in Social Psychology* (pp.49-72). Psychology Press.
Valdez, J. y Salinas, J. (2019). Análisis de las respuestas de estudiantes de bachillerato a problemas sobre la distribución normal. En J. M. Contreras, M. Gea, M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística* (pp. 1-10). Departamento de didáctica de la matemática de la Universidad de Granada.
Valverde, M. (2017). *Un estudio de la presentación de la distribución normal en los textos de bachillerato* [Tesis de Máster, Universidad de Granada].
Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D. y Verschaffel, L. (2003). The illusion of linearity: Expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics, 53,* 113–138.
Vilca, M. (2015). Tipificación de los errores que se presentan al identificar una variable aleatoria de distribución binomial en problemas contextualizados [Tesis de Máster, Pontificia Universidad Católica del Perú].
Zieffler, A., Garfield, J., delMas, R. y Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 7*(2), 40-58.

2.3 Antecedentes complementarios

En esta sección se presentan trabajos que no han sido considerados en la publicación de BEIO (Bizet et al., 2022) y que abordan la enseñanza y/o aprendizaje de variable aleatoria o distribuciones de probabilidad. Para su análisis se utilizó el mismo método empleado en el artículo anterior, aunque se incorporaron nuevos descriptores (distribución binomial y comprensión; distribución normal y comprensión) y extendió el límite de año de publicación (hasta 2022) Así la presentación de los resultados se organiza de igual manera.

2.3.1 Enseñanza de la variable aleatoria y su distribución de probabilidad

En relación con investigaciones sobre aspectos epistemológicos de los temas en cuestión se han incorporado los trabajos de Amrani y Zaki (2015), García-García et al. (2022), Ruiz y Albert (2013) y Stahl (2006).

Amrani y Zaki (2015), se focalizan en analizar el origen y evolución histórica de la variable aleatoria y distribución de probabilidad. Entre sus resultados obtienen que la evolución de la definición de aquellos conceptos está íntimamente relacionada con: i) el desarrollo de los conceptos de función y probabilidad; ii) los problemas que surgían (por ejemplo, expectativas de ganancia) y la naturaleza de su contexto (juegos de azar). Además, advierten que los estudiantes podrían manifestar el error de afirmar que la variable aleatoria es una variable arbitraria o confundirla con su distribución de probabilidad.

Así mismo Ruiz y Albert (2013) exploran el desarrollo histórico de la variable aleatoria, determinando seis fases importante en su evolución: i) utilización intuitiva de la variable estadística, por ejemplo, para realizar estudios estadísticos en un censo o analizar mediciones en astronomía; ii) utilización intuitiva de la variable aleatoria, en juegos de azar o cálculos del error cuando se realizaban mediciones; iii) utilización explícita de la variable aleatoria aunque sin su formalización, en un comienzo es considerada como resultado numérico de una situación aleatoria luego se establece como una variable relacionada con el azar; iv) relación informal entre la variable aleatoria y estadística, son utilizadas tanto herramientas de estadística como de probabilidad en problemas de la sociedad, biología o economía; v) establecimiento de los axiomas de probabilidad, es reconocida la variable aleatoria como una función medible y se formaliza su teoría; iv) establecimiento de la inferencia estadística, es formalizada la relación entre la variable estadística y variable aleatoria.

En tanto García-García et al. (2022) indagan el origen y desarrollo histórico de la distribución binomial, estableciendo como resultado cinco etapas en su evolución, que involucran problemas de complejidad creciente: i) cálculo de combinaciones y casos posibles y favorables de una situación binomial; ii) cálculo de combinaciones a través del triángulo de Pascal y probabilidades mediante el enfoque clásico; iii) identificación de los valores de los parámetros p y q, y determinación de probabilidad utilizando la regla del producto en una situación binomial; iv) cálculo de probabilidades en una situación binomial para el caso particular $p = \frac{1}{2}$; y v) cálculo de probabilidades en una situación binomial para cualquier valor de n, p y q, y determinación de su media y desviación estándar.

Mientras que Stahl (2006) explora la evolución histórica de la distribución normal, identificando como resultado dos etapas fundamentales para su desarrollo, que coinciden cronológicamente: i) la búsqueda de una herramienta para aproximar distribuciones binomiales con un número grande de repeticiones independiente de un experimento Bernoulli, momento que se establece el teorema de Moivre (aproximación de la binomial por la normal); y ii) la búsqueda de una distribución de probabilidad que modelara el comportamiento de errores cometidos al realizar diferentes mediciones, instancia que se obtiene mediante el método de mínimos cuadrados la función de densidad para determinar el error.

Respecto a estudios sobre la didáctica de la variable aleatoria y distribuciones de probabilidad se han integrado los trabajos de Derouet (2019), Garfield et al. (2008), Godino (2014), Salinas-Herrera y Salinas-Hernández (2022) y Stephenson et al. (2009).

Concretamente Derouet (2019) diseña y prueba una secuencia de enseñanza para introducir la noción de función de densidad y la relación entre la probabilidad e integral en grado 12 (17 a 18 años), fundamentada principalmente en la teoría Espacio de Trabajo Matemático (Kuzniak, Tanguay y Elia, 2016). La secuencia está constituida por dos problemas: i) en el primero, se debe simular la situación propuesta y construir un histograma asociado a una muestra mediante uso de software, luego identificar la función de densidad que modela el problema y calcular probabilidades mediante el cálculo de áreas de figural simples; ii) en el segundo, se requiere evaluar las frecuencias obtenidas de los datos entregados y establecer la función de densidad que modela el problema, después analizar las propiedades de aquella función y calcular probabilidades por medio del cálculo de área aproximada (uso de softwares o método de rectángulos). Así concluyen que el éxito de la actividad matemática por parte de los estudiantes, está influenciada tanto por en el diseño de la secuencia, donde se requieren identificar las posibles actuaciones de estos, como por las intervenciones del profesor en la implementación, las que deben ser una guía en el trabajo de aquellos.

En cambio, Godino (2014) diseña una secuencia de tareas que profundizan nociones elementales de probabilidad en estudiantes universitarios (de educación primaria), fundamentada en el EOS (Godino et al., 2007). En la primera tarea se requiere establecer si el juego propuesto es equitativo, el cual consiste en lanzar dos dados de seis caras y sumar los puntos obtenidos, donde un jugador gana si la suma es 6, 7, 8 o 9 y gana el otro si la suma es distinta. La siguiente tarea consiste en simular el juego 100 veces con un software e indagar los datos. En la tercera y cuarta tarea es necesario identificar los objetos matemáticos primarios involucrados en el juego (situaciones-problemas, lenguaje, conceptos, proposiciones, procedimientos y argumentos). Luego se debe determinar la variable aleatoria y distribución de probabilidad asociada al juego para argumentar que un jugador posee ventaja. El autor concluye que la configuración de objetos primarios es una herramienta útil para reconocer aquellos involucrados en la solución de una tarea, identificar posibles dificultades de aprendizaje y decidir estrategias de enseñanza.

En tanto Stephenson et al. (2009) elaboran una secuencia de cuatro actividades para presentar distribuciones de probabilidad de variables aleatorias discretas e ilustrar la inferencia sobre proporciones, utilizando dados Golo que están basados en el juego de golf. En la primera actividad se estudian las características de la distribución binomial, calculan probabilidades y determina su función de distribución (desde el enfoque frecuencia y clásico de probabilidad). En las siguientes actividades se aborda el proceso de toma de muestras repetidas de una población y son empleados los datos para realizar inferencia sobre la proporción desconocida de la población, construir un intervalo de confianza y desarrollar una prueba de hipótesis. Los autores concluyen que el juego de dados Golo es una herramienta interesante de utilizar para estudiar conceptos de estadística y probabilidad, debido a que posee una diversidad de dados y se relaciona con un deporte competitivo como el golf.

Por su parte, Garfield et al. (2008) diseñan una secuencia de enseñanza para un curso introductorio de estadística. Aquella está integrada por tres actividades: i) en la primera, es necesario simular dos situación reales (natalidad de niñas y niños y problema de monte hall) utilizando material manipulativo y un software, luego se debe construir distribuciones de probabilidad para representarlas; ii) en la segunda, se requiere simular situaciones sobre lanzamiento de monedas y dados, y extracción de cartas, empleando material concreto y un software, después se necesita modelarlas mediante distribuciones binomiales; iii) en la tercera, a partir de datos recogidos sobre diferentes variables (altura de un estudiante, amplitud de su mano, etc.) se debe identificar cuáles de estas poseen un conjunto de datos que al graficarlas, mediante software, tienen una distribución normal, además de razonar sobre algunas característica de este modelo de probabilidad.

Más en la actualidad, Salinas-Herrera y Salinas-Hernández (2022) elaboran e implementan un problema para en grado 12 introducir la distribución normal como aproximación de la binomial, utilizando un software y sustentados en el Enfoque Documental de la Didáctica (Trouche et al., 2020). El problema consiste en calcular la probabilidad que una estudiante responda correctamente 120 preguntas de selección múltiple de un total de 200, donde cada una posee tres alternativas. Entre sus resultados observan que debido a que los participantes tenían experiencia previa en trabajar con el software involucrado, algunos propusieron utilizarlo para resolver el problema. También los estudiantes, aunque no lograron realizar por sí solos la simulación de la situación, algunos identificaron que la probabilidad solicitada es baja. Así se concluyen que la simulación con el software empleado no es suficiente para mejorar la comprensión de la distribución normal como aproximación de la binomial.

2.3.2 Aprendizaje de la variable aleatoria y su distribución de probabilidad

En las investigaciones sobre aspectos cognitivos de los temas de interés, se han incorporado en el contexto escolar los trabajos de Begué et al. (2020), Doukhan y Gueudet (2019) y Shin (2012). Mientras que en el nivel universitario se han integrado los estudios de Armah et al. (2016), Derouet et al. (2018), Fernández et al. (2022), Hwang y Chae (2020) y Yuan y Li (2012).

Dentro del ámbito escolar, Doukhan y Gueudet (2019) evalúan la comprensión de los tópicos vinculados a la variable aleatoria fundamentadas en la teoría Antropológica de lo Didáctico (Chevallard, 1999) y la teoría de Registro de Representaciones Semiótica (Duval, 2006). Para ello aplicaron un cuestionario a 44 estudiantes franceses de grado 11 (16 a 17 años), obteniendo como resultado que el 84% de la muestra determina su función de probabilidad y un 25% de ellos la representa en un diagrama sagital. Además el 86,4% de los participantes calcula e interpretan su media. Las autoras concluyen que existe confusión entre la variable aleatoria y su función de probabilidad.

También en dicho contexto, en torno a los modelos probabilísticos, Begué et al. (2020) valoran la comprensión de la media de la distribución binomial en 127 estudiantes españoles de grado 12 (17 a 18 años) a través de la resolución de una tarea. Entre sus resultados evidencia que el 57,5% de la muestra propone una aproximación adecuada de la media, y predomina en las justificaciones correctas la asignación de probabilidad desde el enfoque frecuencial (56,7%). A modo de conclusión los autores señalan que los participantes manifiestan confusión entre suceso aleatorio y suceso equiprobable.

Por su parte, Shin (2012) indaga la compresión de la distribución normal en 92 estudiantes surcoreanos de grado 11 y 12 (16 a 18 años) mediante la resolución de un cuestionario. Los resultados muestran que, en un problema el 88% de los participantes realizan el procedimiento de tipificación para los dos valores de la variable aleatoria continua involucrada y solo el 2,2% calcula la probabilidad solicitada utilizando la propiedad de los intervalos centrales ($\pm 3\sigma$). También el 7,6% de la muestra argumenta que la altura de las personas sigue una distribución normal, debido a que aquellos valores se agrupan alrededor de la media y la cantidad de crecimiento, que difiere mucho de la media, es cada vez más pequeña. Entre las conclusiones se mencionan que existe el error de afirmar que toda variable aleatoria continua se distribuye de forma normal.

En el ámbito universitario Armah et al. (2016) valoran la comprensión de la desviación estándar de la variable aleatoria. Participaron 430 estudiantes ghaneses recién ingresados a la carrera de educación matemáticas, quienes respondieron a un cuestionario. Los resultados constatan que el 8,1% de la muestra calcula su valor y el 3% logra interpretarla. Además, el 30,9% de los estudiantes selecciona correctamente el gráfico que posee mayor desviación estándar y el 5,3% argumenta su elección debido a la mayor dispersión de los datos. De esta manera, se concluye que los participantes no logran comprender el procedimiento de calcular la desviación estándar ni el significado de su valor.

Luego, Derouet et al. (2018) analizan el uso del cálculo integral para resolver tareas sobre distribuciones de probabilidad de variables aleatorias continuas y viceversa, con base en la teoría Antropológica de lo Didáctico (Chevallard, 1999). Para ello aplicaron un cuestionario a 82 estudiantes franceses de primer año de ingeniería evidenciando como resultado que el 39% de sus participantes conoce las propiedades que definen una función de densidad y el 70% identifica la probabilidad de una variable aleatoria continua como una integral.

Respecto al aprendizaje de los modelos de probabilidad en el contexto universitario, Wroughton y Cole (2013) evalúan la diferenciación entre la distribución binomial, hipergeométrica y binomial negativa. Participaron 20 estudiantes estadounidenses los cuales respondieron un pretest y un post test después de un proceso de instrucción sobre aquellos temas.

Los resultados evidencian, en ambas versiones del test, que la mayoría de las respuestas incorrectas de los participantes se presentan cuando deben reconocer situaciones binomiales.

Posteriormente, Fernández et al. (2022) exploran la comprensión de la distribución binomial fundamentados en elementos del EOS (Godino et al., 2007). Para tal efecto, aplicaron un problema a 22 futuros profesores chilenos de matemáticas, después de un proceso de instrucción sobre el tópico en cuestión. Los resultados demuestran que: el 91% identifica la situación binomial propuesta; el 82% determina los parámetros de la binomial; 73% emplea la función de probabilidad de la binomial para determinar la probabilidad solicitada; y el 64% calcula su media. También se observó que alrededor del 90% de los participantes emplean el lenguaje verbal y simbólico vinculado a la distribución binomial y existe ausencia de la argumentación mediante el razonamiento verbal-deductivo o por medio de representaciones (lenguaje gráfico y tabular).

En tanto que Yuan y Li (2012) indagan el conocimiento de la distribución normal fundamentados en el Conocimiento Tecnológico Pedagógico del Contenido (Niess, 2005). Para ello, primero implementan en 22 futuros profesores chinos de matemáticas un curso de desarrollo docente, que incluye la enseñanza de la normal empleando tecnología. Después los participantes diseñan e implementan en grado 11(17 a 18 años) una clase sobre el tema. Los resultados constatan que el 23,1% de los futuros profesores proponen situaciones reales (estatura y puntajes) para ejemplificar la normal y utilizan softwares (Fathom o Minitab) para graficarla y mostrar sus áreas sombreadas.

Finalmente, Hwang y Chae (2020) analizan el conocimiento de la distribución normal basados en el modelo de Conocimiento Matemático para la Enseñanza (Ball et al., 2008). Participaron 24 futuros profesores surcoreanos de matemáticas que habían cursado previamente una asignatura sobre tópicos estocásticos, quienes respondieron a un cuestionario. En los resultados se observa que: el 54,2% tiene experiencia en enseñar estadística; el 79,2% identifica la variable aleatoria continua involucrada en el problema y realiza el procedimiento de tipificación; y el 62,5% reconoce que la distribución binomial se puede aproximar a la normal, pero no explicita las condiciones que deben cumplir los parámetros de la binomial (p, q y n).

CAPÍTULO 3: ANÁLISIS DE LIBROS DE TEXTO Y NORMATIVA CURRICULAR CHILENA

3.1 Introducción

3.2 **Estudio 2.** Situaciones-problemas sobre variable aleatoria y sus aplicaciones en de distribuciones probabilidad según libros de texto chilenos

3.2.1 Introducción

3.2.2 Marco teórico

3.2.3 Metodología

3.2.3.1 Muestra y unidades de análisis

3.2.3.2 Procedimiento de análisis

3.2.3.3 Guía de situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad

3.2.4 Resultados

3.2.4.1 Distribución de las actividades de aprendizaje

3.2.4.2 Campos de problemas

Campo de problema 1: Identificar la variable aleatoria como una función presente en el contexto probabilístico

Campo de problema 2: Reconocer la distribución de probabilidad de una variable aleatoria discreta como herramienta que permite ver la variación aleatoria

Campo de problema 3: Reconocer la distribución de probabilidad de una variable aleatoria continua como una herramienta que permite ver la variación aleatoria

Campo de problema 4: Establecer la función de distribución de una variable aleatoria discreta

Campo de problema 5: Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria

Campo de problema 6: La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real

Campo de problema 7: La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real

Campo de problema 8: Aproximar distribuciones de variables aleatorias discretas con un número grande de valores

3.2.4.3 Contexto de las actividades

3.2.5 Conclusiones

3.3 **Estudio 3.** Objetos matemáticos ligados a la variable aleatoria y sus aplicaciones sobre distribuciones de probabilidad en libros de texto de educación secundaria chilena

3.3.1 Introducción

3.3.2 Marco teórico

3.3.3 Metodología

3.3.4 Resultados

3.3.4.1 Etapa 1: Objetos matemáticos presentes en el currículo escolar

3.3.4.2 Etapa 2.1: Unidades y lecciones indagadas en libros de texto

3.3.4.3 Etapa 2.2: Objetos matemáticos presentes en libros de texto

Lenguaje

Conceptos

Proposiciones

Procedimientos

Argumentos

3.3.4.4 Fase 3: Contraste entre currículo escolar chileno y libros de texto

3.3.5 Discusión y conclusiones

3.1 Introducción

En el presente capítulo se lleva a cabo el análisis de la normativa curricular chilena vigente durante el desarrollo de esta tesis doctoral y la indagación de libros de texto chilenos dirigidos a los grados 10 a 12 (15 a 18 años), centrándose en la identificación de los objetos matemáticos primarios ligados a la variable aleatoria y distribuciones binomial y normal. Las situaciones-problemas sobre los temas de interés que son reconocidas en el currículo escolar y libros de texto chilenos, se abordan en el apartado 3.2 mediante el Estudio 2. Situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad según libros de texto chilenos, publicado en Profesorado, revista de currículum y formación de profesorado. También en este estudio son identificados los contextos de aquellas situaciones-problemas propuestas en los textos indagados.

El lenguaje, los conceptos, las proposiciones, los procedimientos y los argumentos vinculados a los temas en cuestión, que son reconocidos en los lineamientos curriculares y libros escolares chilenos, se presentan en el apartado 3.3 a través del Estudio 3. Objetos matemáticos ligados a la variable aleatoria y sus aplicaciones sobre distribuciones de probabilidad en libros de texto de educación secundaria chilena, publicado en PNA. Así los objetos matemáticos primarios en torno a la variable aleatoria y distribuciones binomial y normal determinados a partir del análisis de la normativa curricular chilena permiten constituir el significado institucional de referencia de la presente tesis doctoral. Mientras que los objetos matemáticos primarios establecidos desde la indagación de libros de texto chilenos componen el significado institucional pretendido de esta investigación.

3.2 Estudio 2[2]. Situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad según libros de texto chilenos

Bizet, V., Molina-Portillo, E. y Contreras, J.M. (2023a). Situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad según libros de texto chilenos. *Profesorado, revista de currículum y formación de profesorado*, *27*(2), 351-382.

[2] Para evitar confusión y repetición en la numeración, se ha modificado la relativa a los epígrafes pertenecientes a los estudios que forman parte del compendio, haciendo referencia al capítulo en que se ubica cada estudio. También se ha cambiado la numeración de Tablas y Figuras con respecto a la versión publicada del artículo, haciendo referencia al estudio en que se enmarcan.

Resumen

Este articulo analiza las situaciones-problemas ligadas a la variable aleatoria y sus aplicaciones sobre las distribuciones de probabilidad en cinco libros de texto chilenos de educación secundaria, fundamentado en herramientas del Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos. Aquel objeto matemático es analizado con base a una guía de situaciones-problemas de los tópicos en cuestión diseñada a partir del currículo chileno oficial. Los resultados muestran que en los libros analizados existe ausencia y mínima presencia de actividades sobre calcular probabilidades asociadas a distribuciones binomial y normal empleando una herramienta tecnológica, y diferenciar entre variables aleatorias y variables con dependencia funcional, respectivamente. Aquellos resultados demuestran falta de alineación entre la propuesta del Ministerio de Educación chileno y los textos analizados. Esta investigación proporciona información valiosa a los profesores que enseñan la variable aleatoria y/o distribuciones binomial y normal en la escuela, investigadores interesados en desarrollar estudios sobre su comprensión y quienes diseñan libros de texto escolares para mejorar la presentación de estos temas.

Palabras clave: distribuciones de probabilidad, Enfoque Ontosemiótico, libro de texto, objetos matemáticos, variable aleatoria.

Abstract

This article analyzes the problems-situations related to the random variable and their applications about probability distributions in five Chilean textbooks of high school, based on the Onto-Semiotic Approach to Mathematical. That mathematical object is analyzed based on a guide of problems-situations of the topics in question designed from the official Chilean curriculum. The results show that in the textbooks analyzed are absence and minimum presence of activities about calculate probabilities associated with binomial distributions and normal distributions using a technology tool, and differentiate between random variables and variables with functional dependency, respectively. Those results demonstrate lack of alignment between the proposal of the Chilean Ministry of Education and textbooks analyzed. This research provides valuable information for the professor that teaches the random variable and/or binomial and normal distributions at the school, the researcher who is interested in developing studies about its understanding and whom designs school textbooks to improve the presentation of this theme.

Keywords: probability distribution, Onto-Semiotic Approach, textbooks, mathematical objects, random variable.

3.2.1 Introducción

Dentro de la enseñanza estocástica (estadística y probabilidad) en la escuela, el razonamiento en torno a la variable aleatoria es fundamental, debido a que es un tema que posee un rol importante en la evolución del cálculo de probabilidades y se involucrada en diversas situaciones aleatoria de la vida real (Batanero, 2004; Heitele,1975).También las distribuciones de probabilidad de variables aleatorias es un tópico importante, pues permite vincular la estadística con experiencias del mundo real, enseñar la estructura de esta disciplina y lograr una comprensión profunda de aquellas distribuciones a medida que los estudiantes maduran su conocimiento sobre estadística (Burrill y Biehler, 2011).

Este escenario motivó que diversos países incluyan en su currículo escolar los temas en cuestión. Según la denominación K-12 (Franklin et al., 2005), kindergarden más 12 grados consecutivos, Estados Unidos (National Council of Teachers of Mathematics [NCTM], 2000; National Governors Association Center for Best Practices [NGACBP] y Council of Chief State School Officers [CCSSO], 2010) abordada la variable aleatoria entre los grados 9 y 12 (14 a 18 años), e introduce su función de probabilidad mediante los enfoques de probabilidad frecuencial y clásico. En tanto Chile (Ministerio de Educación [MINEDUC], 2015) inicia el estudio de la variable aleatoria en grado 10 (15 a 16 años), contexto donde trabaja la función señalada por medio del enfoque clásico. Mientras que España (Ministerio de Educación, Cultura y Deporte [MECD], 2015) introduce el tratamiento de la variable aleatoria entre los grados 11 y 12 (16 a 18 años), además de estudiar su función de probabilidad desde el enfoque de probabilidad clásico.

Por su parte, las distribuciones binomial y normal, en Estados Unidos (NCTM, 2000; NGACBP y CCSSO, 2010) son abordadas entre los grados 9 y 12 (14 a 18 años), a través de los enfoques de probabilidad frecuencial y clásico, allí es utilizada la tecnología digital para simular experimentos aleatorios y aproximar las distribuciones de frecuencias a las de probabilidad. En Chile (MINEDUC, 2019a) durante grado 12 (17 a 18 años) y España (MECD, 2015) entre los grados 11 y 12 (16 a 18 años), dichas temáticas son trabajadas bajo los mismos enfoques y metodología descrita.

En esta perspectiva, los libros de texto escolares son instrumentos valiosos que permiten que el currículo dispuesto por un organismo gubernamental se ejecute en la educación escolar de un país mediante una presentación didáctica (Occeli y Valeiras, 2013). Sin embargo, se presenta el problema que existen escasos estudios sobre el tratamiento de la variable aleatoria o distribuciones de probabilidad en aquel recurso didáctico.

Respecto a la variable aleatoria, Ortiz (2002) caracterizó su significado institucional desde el análisis de 11 libros de texto españoles para primer año de bachillerato (grado 11). Entre sus resultados propone nueve situaciones-problemas diferentes vinculadas al concepto y obtiene que solo un libro presenta mayor diversidad con cuatro de estas situaciones y en tres libros existe ausencia de ellas. Por su parte, Doukhan y Gueudet (2019) estudiaron su tratamiento en libros franceses, por medio de un análisis comparativo entre dos textos, uno dirigido a grado 11 y otro a grado 12. Como resultado obtienen que los documentos proponen tipos de problemas similares, con dos contextos diferentes, situaciones matemáticas-teóricas (sin contexto) y no-matemáticas (con contexto: vida real, artificial y otras ciencias), predominando este último en ambos.

En relación con las distribuciones de probabilidad, Valverde (2017) determinó el significado institucional de la normal desde el estudio a dos libros de texto españoles de bachillerato (grado 11 y 12). En cuanto a la variedad de situaciones-problemas propuestas, destaca que solo un libro analizado presenta cuatro tipos diferentes, a diferencia del otro texto que propone solo dos de ellas. Además, el contexto de aquellos problemas es variado (social, científico, personal, descontextualizado y laboral, predominando este último). En cambio, sobre el tratamiento de la binomial en textos escolares, no se encontraron estudios.

De esta manera, es necesario desarrollar investigaciones que analicen el tratamiento otorgado a la variable aleatoria y sus aplicaciones sobre las distribuciones de probabilidad en libros de texto escolares, e informen ¿cómo son abordados aquellos tópicos probabilísticos en tales instrumentos de enseñanza? Por tanto, el presente trabajo, fundamentado en el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos, tiene como objetivo analizar los objetos matemáticos ligados a la variable aleatoria y sus aplicaciones relativas a las distribuciones de probabilidad en cinco libros de texto chilenos de educación secundaria (grado 9 a 12).

3.2.2 Marco teórico

Este estudio se fundamenta en la herramienta de Configuración Ontosemiótica del Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos (Font et al., 2013). En ella el origen del conocimiento matemático está en la resolución de situaciones-problemas, para las cuales existen determinadas prácticas (actuaciones o expresiones verbal, gráfica, etc.) realizadas por una persona o compartidas en una institución, donde intervienen y emergen diversos objetos matemáticos.

Para el EOS un objeto matemático es "*cualquier entidad material o inmaterial que participa en la práctica matemática apoyando y regulando su realización"* (Godino et al., 2019, p. 40). En este contexto se propone una tipología de objetos primarios (Godino et al., 2007):

- Situaciones-problemas: tareas, ejercicios o problemas que incitan a realizar una actividad matemática.
- Lenguaje: términos matemáticos en sus diversos registros de representación (verbal, simbólico, tabular y gráfico), utilizados para presentar un enunciado o resolver una situación-problema.
- Conceptos: descripciones o definiciones aludidas, explícita o implícitamente, al resolver un problema.
- Proposiciones: enunciados sobre características o propiedades de objetos matemáticos aplicados en la resolución de una situación-problema.
- Procedimientos: operaciones, técnicas de cálculo, procedimientos o algoritmos desarrollados para resolver un problema
- Argumentos: enunciados que un sujeto emplea en la explicación o validación de procedimientos o solución de una situación-problema.

Así, un conjunto de situaciones-problema vinculadas recíprocamente, que comparten sus representaciones, procesos o soluciones similares constituyen un campo de problema (Godino, 1999). Además, la interacción de los objetos matemáticos primarios establecidos en la solución a un problema o en el proceso de enseñanza-aprendizaje de un tema, genera configuraciones de carácter: cognitivo, referida al conocimiento sobre el razonamiento de estudiantes, sus dificultades y errores; epistémico, relativa al conocimiento matemático a enseñar desde la perspectiva institucional (Font et al., 2013). Esta última noción es de utilidad para realizar una indagación y descripción de las características de libros de texto matemáticos en torno a un tópico (Font y Godino, 2006).

Por consiguiente, en la presente investigación fue utilizada la configuración ontosemiótica epistémica para identificar y analizar la tipología de objetos matemáticos que intervienen en los libros de texto escolares, cuando es tratada la variable aleatoria y sus aplicaciones sobre las distribuciones de probabilidad, centrándose en el análisis del principal objeto primario, las situaciones-problemas. En esta perspectiva, las tareas, ejercicios o problemas relacionados con la tecnología permiten valorar el conocimiento avanzado de un tema pues, aunque esta herramienta pudiera usarse en la escuela entre los grados 11 y 12 (16 a 18 años), en la actualidad no es frecuente su uso (Gea, 2014).

3.2.3 Metodología

Esta investigación se desarrolló bajo un enfoque cualitativo y es de tipo exploratoria-descriptiva (Hernández at al., 2014). En ella se empleó el análisis de contenido (Krippendorff, 1990), técnica distintiva para indagar sistemática y exhaustivamente documentos escritos con el propósito de producir generalizaciones (Zapico, 2006).

3.2.3.1 Muestra y unidades de análisis

En Chile existen tres editoriales que diseñan libros de texto de Matemáticas para educación secundaria (grados 9 a 12): Editorial Santillana, Ediciones SM y Editorial crecer pensando. En la educación pública, para cada grado educativo, una de las editoriales se adjudica el proyecto concursable con vigencia de 4 años. Mientras que en la educación privada aquellas editoriales ofrecen sus proyectos, en ocasiones más de uno para cada nivel educativo, donde la más utilizada es la Editorial Santillana seguida de Ediciones SM. En el periodo de desarrollo de esta investigación (año 2021) para los grados 10 a 12, donde se enseña los temas de interés: en el ámbito público Ediciones SM tenía el proyecto adjudicado; y en el contexto privado la Editorial Santilla ofrecía el proyecto Bicentenario, Ediciones SM proponía el proyecto Sabia y la Editorial crecer pensando presentaba un proyecto sin nombre.

Así en la presente investigación, la elección de la muestra fue intencionada y estuvo constituida por dos series de libros de texto chilenos. Como muestra la Tabla 2.1, cada una contempló los grados 10 (15 a 16 años), 11 (16 a 17 años) y 12 (17 a 18 años), y los textos que las integraron permanecían vigentes para el curso escolar 2021. Las dos series son de prestigiosas editoriales del país, la primera de ellas correspondió a los libros empleados tanto en la educación pública como subvencionada y entregados gratuitamente por el MINEDUC, mientras que la segunda serie incluyó los textos más utilizados en la enseñanza privada.

Tabla 2.1

Libros de texto utilizados en el análisis

Serie	Código	Nivel educativo	Título	Editorial	Año
1	T1	10°	Texto del estudiante de matemática 2° medio	Ediciones SM	2020
	T2	11° y 12°	Texto del estudiante de matemática 3°y 4° medio		2019
2	T3	10°	Matemática 2 proyecto bicentenario	Santillana	2009
	T4	11°	Matemática 3 proyecto bicentenario		2009
	T5	12°	Matemática 4 proyecto bicentenario		2014

Nota. Elaboración propia.

En los cinco libros se analizó las lecciones sobre los temas variable aleatoria, distribución binomial y/o distribución normal. Así las unidades de análisis definidas correspondieron a las actividades propuestas en cada texto que incitan en los estudiantes el aprendizaje de los tópicos en cuestión.

3.2.3.2 Procedimiento de análisis

El estudio a las lecciones señaladas fue realizado mediante un proceso cíclico e inductivo, como se recomienda para el análisis de datos de tipo cualitativo (Gil, 1994). Aquello posibilitó la codificación de datos y construcción de tablas de frecuencia y porcentaje, para conseguir resultados respecto a los objetos matemáticos primarios que intervienen en el tratamiento de los conceptos probabilísticos de interés. Este fue organizado consideraron las siguientes etapas:

1 Identificación y selección de las lecciones vinculadas a los tópicos variable aleatoria y distribuciones de probabilidad.

2 Clasificación de las actividades presentes, reconociendo el o los tipos de situaciones-problemas involucrados en cada una. Para ello fue utilizada la Guía de Situaciones-Problemas sobre Variable Aleatoria y sus aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar Chileno (GSP-VADP: Bizet et al., 2023), que caracteriza el

significado institucional de referencia de los tópicos en cuestión, inferido desde el currículo chileno oficial y lineamientos internacionales.

3 Categorización de las actividades según el contexto empleado en cada una.

4 Elaboración de tablas y gráficos que resumen la información obtenida del análisis y permiten desarrollar un análisis descriptivo.

3.2.3.3 Guía de situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad

El proceso de construcción de la GSP-VADP[3] (Bizet et al., 2023) está fundamentado en el modelo de Vásquez y Alsina (2015) y se organizó en dos etapas que se describen a continuación.

Etapa 1, identificación de campos de problemas sobre los temas de interés a partir de un análisis epistemológico. Se han indagado investigaciones sobre aspectos epistemológicos de la variable aleatoria (Amrani y Zaki, 2015; Dinges, 2005; Ruiz, 2013), distribución binomial (García-García et al., 2022; Vilca, 2015) y distribución normal (Stahl 2006; Tauber, 2001), las cuales se caracterizan por indagar su origen, desarrollo histórico y el estudio de cada concepto desde la matemática. Por medio de aquel análisis, en un proceso cíclico e inductivo, se establecieron ocho campos de problemas (C-P) ligados a los temas en cuestión.

Etapa 2, reconocimiento de situaciones-problemas sobre el tema de interés desde un análisis de lineamientos curriculares. En un primer momento fue indagado el currículo escolar chileno (MINEDUC, 2015; MINEDUC, 2016; MINEDUC, 2019a; MINEDUC, 2019b; MINEDUC 2019c) entre los grados 10 y 12. Debido a que su construcción se sustentan en diversas fuentes, una de ellas la revisión de lineamientos curriculares internacionales principalmente de países pertenecientes a la Organización para la Cooperación y el Desarrollo Económico (MINEDUC, 2009), en un segundo momento se estudiaron documentos curriculares americanos (NCTM, 2000; NGACBP y CCSSO, 2010) dirigido a los grados 9 a 12, y el conocido proyecto GAISE (Franklin et al., 2005).

Así mediante un análisis de contenido a los documentos señalados se identificaron normativas vigentes (unidades de análisis) en torno a la enseñanza-aprendizaje de la variable aleatoria y distribuciones de probabilidad, luego estas fueron clasificadas en los ocho campos de

[3] El proceso de construcción de la GSP-VADP se expone en detalle en el capítulo 4, sección 4.2.

problemas (categorías de análisis) previamente definidos. Posteriormente las unidades de análisis se compararon y redujeron debido a que algunas unidades estaban contenidas en otra o no entregaban nueva información, todo ello con el propósito de lograr representar su información en una unidad de análisis final. A partir de estas últimas se procedió a inferir situaciones-problemas (S-P) alusivas a la comprensión de la variable aleatoria y distribuciones de probabilidad, obteniéndose la GSP-VADP.

3.2.4 Resultados

3.2.4.1 Distribución de las actividades de aprendizaje

La Tabla 2.2 muestran la cantidad de actividades asociadas a la variable aleatoria y sus aplicaciones en distribuciones de probabilidad que fueron analizadas en cinco libros de texto chilenos, organizadas según editorial y nivel educativo. Estas fueron 220, observándose mayor presencia en documentos de la editorial Santillana (149) que SM (71).

Tabla 2.2

Frecuencia y porcentaje de actividades analizadas según serie y nivel educativo

Nivel educativo	Serie 1 SM	Serie 2 Santillana	Total
10°	30(54, 55%)	25(45,45%)	55(25%)
11°		53(100%)	53(24,09%)
12°	41(36,61%)	71(63,39%)	112(50,91%)
Total	71(32,27%)	149(67,73%)	220(100%)

Nota. Elaboración propia.

Según muestra la Tabla 2.2, la mayor cantidad de actividades se presentaron en grado 12 (50,91%) y grado 10 (25%), en el primer nivel señalado destaca la editorial Santillana (63,39%) mientras que en el segundo SM (54,55%). En grado 11, Ediciones SM presenta ausencia de actividades sobre los tópicos en cuestión, ello puede deberse que en aquel nivel el currículo vigente no propone explícitamente trabajar con la variable aleatoria y distribuciones de probabilidad.

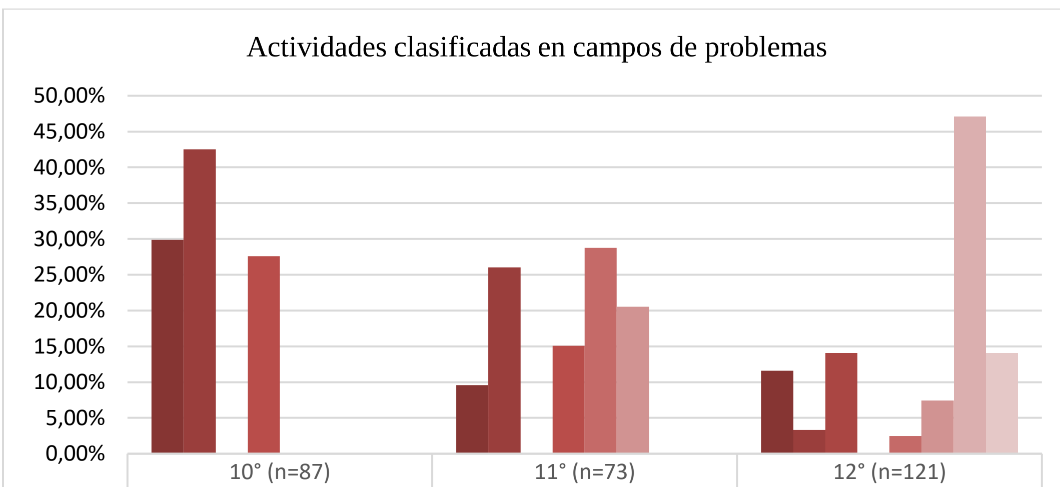
3.2.4.2 Campos de problemas

La clasificación teórica de las situaciones-problemas que otorgan sentido a la variable aleatoria y sus aplicaciones sobre las distribuciones de probabilidad en textos escolares de Chile, se fundamenta en la GSP-VADP (Bizet et al., 2023). Aquella está constituida por ocho campos de problemas, cada uno subdividido en entre dos y ocho situaciones-problemas, evidenciadas en los libros indagados. Cabe mencionar que una actividad presente en los documentos examinados abordó uno o más tipos de situaciones-problemas, es decir, una actividad ha sido contabilizada una o más veces, según corresponda.

La Figura 2.1, muestra la distribución de actividades sobre los campos de problemas presentes en los libros analizados, según nivel educativo. Ella, evidencia que en grado 10 predominaron las actividades: C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico (29,89%); C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria (42,53%); C-P_4 Establecer la función de distribución de una variable aleatoria discreta (27,58%). Resultados que están en consonancia con las directrices curriculares chilenas vigentes, pues en grado 10 se introduce el estudio de la variable aleatoria (discreta) y sus funciones de probabilidad y distribución.

Figura 2.1

Distribución de actividades sobre campos de problemas presentes en los textos analizados por grado



	10° (n=87)	11° (n=73)	12° (n=121)
C-P1	29,89%	9,59%	11,57%
C-P2	42,53%	26,03%	3,31%
C-P3	0%	0%	14,05%
C-P4	27,58%	15,07%	0%
C-P5	0,00%	28,76%	2,48%
C-P6	0%	20,55%	7,44%
C-P7	0%	0%	47,10%
C-P8	0%	0%	14,05%

Nota. Elaboración propia.

En grado 11 se aprecia que prevalecieron las actividades: C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria (26,03%); C-P_5 Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria (28,76%); C-P_6 La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real (20,55%). Respecto a este último campo de problema, los resultados obtenidos no concuerdan con el currículo chileno actual, dado que el trabajo con la binomial es propuesto en grado 12.

También la Figura 2.1, expone que en grado 12 destacaron las actividades: C-P_3 Reconocer la distribución de probabilidad de una variable aleatoria continua como una herramienta que permite ver la variación aleatoria (14,05%); C-P_7 La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real (47,1%); C-P_8 Aproximar distribuciones de variables aleatorias discretas con un número grande de valores (14,05%). Estos resultados son coherentes con los lineamientos curriculares vigentes, debido a que el estudio de la variable aleatoria (continua) y sus aplicaciones en la distribución normal está establecido para grado 12.

A continuación, se describen y ejemplifican las divisiones de los ocho campos de problemas. Las situaciones-problemas que poseen un asterisco (*) son aquellas que han sido observadas en los libros de texto analizados, pero no fueron sugeridas en el currículo chileno.

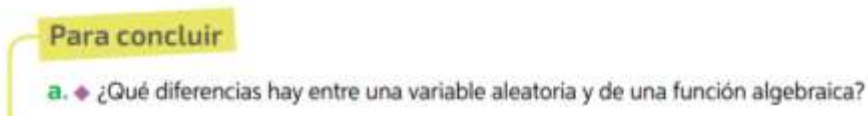
Campo de problema 1: Identificar la variable aleatoria como una función presente en el contexto probabilístico

Reconocer la variable aleatoria ligada a un fenómeno aleatorio es la solución a variados problemas probabilísticos, pero también es el primer paso para estudiar su distribución de probabilidad. Así este se divide en las siguientes cinco situaciones-problemas:

S-$P_{1.1}$ Diferenciar entre variables aleatorias y variables con dependencia funcional. Se observa una actividad que pregunta explícitamente la diferencia entre ambas variables. Las primeras se originan de un fenómeno o experimento aleatorio (aquel indica la posibilidad de que en idénticas condiciones puedan producirse resultados diferentes), mientras que las segundas surgen de un fenómeno determinista (aquel que siempre se produce en igual forma cuando se dan las mismas condiciones). La Figura 2.2 ejemplifica la presente situación-problema.

Figura 2.2

Ejemplo de situación-problema 1.1 (T1)


Para concluir

a. ◆ ¿Qué diferencias hay entre una variable aleatoria y de una función algebraica?

Nota. Extraída de Díaz et al. (2020, p.138).

S-P$_{1.2}$ *Definir variables aleatorias finitas en experimentos aleatorios.* Se exponen tareas que consisten en observar una característica especifica de un experimento aleatorio y determinar en lenguaje verbal una regla de correspondencia que relaciona cada elemento de su espacio muestral con un número real, como se ejemplifica en la Figura 2.3.

Figura 2.3

Ejemplo de situación-problema 1.2, 1.3 y 1.4 (T1)


3. Anota el espacio muestral de los siguientes experimentos aleatorios. Luego, identifica la variable aleatoria involucrada y su recorrido.

a. Lanzar el dado y anotar la cantidad de divisores que tiene el número obtenido.

c. Lanzar las dos fichas de la imagen y anotar la suma de sus resultados.

Nota. Extraída de Díaz et al. (2020, p.138)

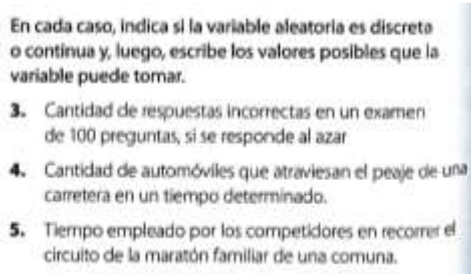
S-P$_{1.3}$ *Identificar dominio de una variable aleatoria finita*. Se presentan actividades que solicitan reconocer cada uno de los elementos del espacio muestral vinculado a un experimento y variable aleatoria propuesta, o relacionado con un experimento aleatorio y una de sus características establecida (ver Figura 2.3).

S-P$_{1.4}$ *Identificar recorrido o valores de una variable aleatoria finita.* Se observan tareas que, a partir de un experimento aleatorio y variable aleatoria propuesta explícita o implícitamente, se requiere determinar el conjunto de valores reales asociado a los elementos del espacio muestral de aquel experimento, la Figura 2.3 ejemplificada aquello.

**S-P*$_{1.5}$ *Diferenciar entre variables aleatorias discretas y continuas*. Se exponen actividades que solicitan ejemplos de variables aleatorias discretas y continuas, o requieren distinguir de un conjunto de variables aleatorias aquellas que son de tipo discretas y continuas, explicitando su recorrido. La Figura 2.4 ejemplifica esta situación-problema.

Figura 2.4

Ejemplo de situación-problema 1.5 (T5)


En cada caso, indica si la variable aleatoria es discreta o continua y, luego, escribe los valores posibles que la variable puede tomar.

3. Cantidad de respuestas incorrectas en un examen de 100 preguntas, si se responde al azar
4. Cantidad de automóviles que atraviesan el peaje de una carretera en un tiempo determinado.
5. Tiempo empleado por los competidores en recorrer el circuito de la maratón familiar de una comuna.

Nota. Extraída de Departamento de Investigaciones Educativas (2014, p.276).

La Tabla 2.3 resume el análisis realizado a los textos sobre el primer campo de problema. Ella evidencia predominio de actividades sobre la S-P$_{1.2}$ definir variables aleatorias finitas (20,19%) y S-P$_{1.4}$ reconocer su recorrido (44,23%), principalmente en grado 10 con un 57,14% y 63,04% respectivamente. También prevalecen las tareas correspondientes a la *S-P$_{1.5}$ diferenciar entre variables aleatorias de tipo discretas y continuas (20,19%), particularmente en grado 12 (66,67%), evidenciándose desarmonía con el currículo chileno actual, pues en este documento no es abordada aquella situación-problema.

Tabla 2.3

Frecuencias y porcentajes sobre el primer campo de problema en textos chilenos

Situaciones-Problemas	10°(n=26)	11°(n=7)	12°(n=14)	Total(n=47)
S-P1.1	1(100%)			1(0,96%)
S-P1.2	12(57,14%)	1(4,76%)	8(38,1%)	21(20,19%)
S-P1.3	12(80%)	2(13,33%)	1(6,67%)	15(14,42%)
S-P1.4	29(63,04%)	8(17,39%)	9(19,57%)	46(44,23%)
*S-P1.5	7(33,33%)		14(66,67%)	21(20,19%)

Nota. Elaboración propia.

Campo de problema 2: Reconocer la distribución de probabilidad de una variable aleatoria discreta como herramienta que permite ver la variación aleatoria

El segundo paso de interés en el estudio de los tópicos en cuestión es trabajar con la función de probabilidad, en este momento de una variable aleatoria discreta, que permite cuantificar la posibilidad de ocurrencia que aquella variable tome cada uno de sus valores. Por tanto, este campo de problema se descompone en las siguientes seis situaciones-problemas:

S-P$_{2.1}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial. Se observan actividades que presentan un experimento aleatorio, solicitando realizar simulaciones de este para recopilar datos y por medio de ellos calcular la probabilidad frecuencial de alguno o todos los valores de la variable aleatoria involucrada, como se ejemplifica en la Figura 2.5, pregunta 3.

Figura 2.5

Ejemplo de situación-problema 2.1 (T4)

Considera el experimento lanzar un dado y extraer una ficha de dominó. Se define la v.a. X: suma de puntos en fichas y dados.

1. Determina los elementos de la v.a. **X**.
2. Genera mediante una planilla de cálculo **100, 5.000** y **10.000** simulaciones del experimento.
3. Obtén la función de probabilidad de la v.a. **X**.

Nota. Extraída de Blanco et al. (2009, p.151).

S-$P_{2.2}$ *Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico.* Se presentan tareas en las que a partir de una función de probabilidad (lenguaje gráfico o tabular) o un experimento aleatorio y variable aleatoria, se requiere calcular la probabilidad de un valor de esta variable por medio de la regla de Laplace. La Figura 2.6, pregunta c, ejemplifica esta situación-problema.

Figura 2.6

Ejemplo de situación-problema 2.2, 2.3 y 2.5 (T1)

5. Considera el experimento aleatorio "lanzar dos dados" y la variable aleatoria *X*: cantidad de números primos que aparecieron en ambos dados.
   a. Escribe el recorrido de la variable aleatoria.
   b. Calcula la función de probablidad de la variable aleatoria.
   c. Calcula $P(X = 1)$ y $P(X < 2)$
   d. Grafica la función de probabilidad.

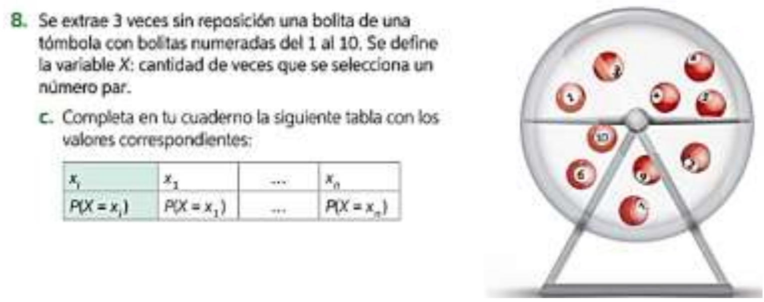
Nota. Extraída de Díaz et al. (2020, p.147).

S-$P_{2.3}$ *Definir la función de probabilidad de una variable aleatoria utilizando la terminología P(X=x).* Se exponen actividades que, por medio de un experimento aleatorio y variable aleatoria discreta, función de probabilidad (lenguaje tabular) o función de distribución (lenguaje gráfico o tabular) de una variable aleatoria discreta, es solicitado identificar los valores de aquella variable y la probabilidad asociada a cada uno. La situación-problema es ejemplificada en la Figura 2.6, pegunta b.

S-P2.4 Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta. Se observan tareas que requieren elaborar una tabla de doble entrada para establecer la probabilidad de cada uno de los valores de una variable aleatoria discreta, o es necesario completar la tabla dada con dicha información (ver Figura 2.7).

Figura 2.7

Ejemplo de situación-problema 2.4 (T1)



8. Se extrae 3 veces sin reposición una bolita de una tómbola con bolitas numeradas del 1 al 10. Se define la variable X: cantidad de veces que se selecciona un número par.

c. Completa en tu cuaderno la siguiente tabla con los valores correspondientes:

x_i	x_1	...	x_n
$P(X = x_i)$	$P(X = x_1)$	...	$P(X = x_n)$

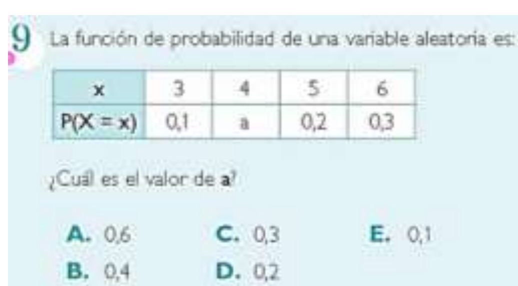
Nota. *Extraída de Díaz et al. (2020, p.142).*

S-P$_{2.5}$ *Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta*. Se exponen tareas que, a través de un experimento aleatorio y variable aleatoria discreta, es necesario construir un gráfico de barras que represente la función de probabilidad vinculada a esta, como es ejemplificado en Figura 2.6, pregunta d.

**S-P*$_{2.6}$ *Calcular el valor de incógnitas para que la función propuesta satisfaga las condiciones de una función de probabilidad*. Se presentan actividades que proponen la función de probabilidad de una variable aleatoria discreta y solicitan determinar el valor de incógnitas correspondientes a la probabilidad de ciertos valores de la variable involucrada, de manera que la función dada cumpla las condiciones de una función de probabilidad. La Figura 2.8 ejemplifica esta situación-problema.

Figura 2.8

Ejemplo de situación-problema 2.6 (T3)



9 La función de probabilidad de una variable aleatoria es:

x	3	4	5	6
P(X = x)	0,1	a	0,2	0,3

¿Cuál es el valor de **a**?

A. 0,6 **C.** 0,3 **E.** 0,1
B. 0,4 **D.** 0,2

Nota. *Extraída de Blanco et al. (2009, p.333).*

La Tabla 2.4 resume el análisis desarrollado a partir de los libros sobre el segundo campo de problema. Aquella muestra que destacan las actividades sobre la S-P$_{2.3}$ definir la función de probabilidad de una variable aleatoria (35,92%) y S-$P_{2.2}$ calcular probabilidades asociadas a los valores de aquella variable desde el enfoque clásico de probabilidad (27,18%), principalmente en grado 10 con un 75,67% y 57,14% respectivamente. Los resultados obtenidos están en consonancia con el currículo chileno actual, dado que este promueve en grado 10 el trabajo con la función de probabilidad bajo el enfoque clásico. Por otra parte, destaca la presencia, en pequeña proporción, de tareas sobre la *S-$P_{2.6}$ calcular el valor de incógnitas de manera que la función propuesta sea de probabilidad (11,65%), la cual no ha sido abordada en los lineamientos curriculares chilenos vigentes.

Tabla 2.4

Frecuencias y porcentajes sobre el segundo campo de problema en textos chilenos

Situaciones-Problemas	10°(n=37)	11°(n=19)	12°(n=4)	Total(n=60)
S-P2.1	1(33,33%)	2(66,67%)		3(2,91%)
S-P2.2	16(57,14%)	9(32,14%)	3(10,71%)	28(27,18%)
S-P2.3	28(75,67%)	7(18,92%)	2(5,41%)	37(35,92%)
S-P2.4	3(100%)			3(2,91%)
S-P2.5	13(65%)	5(25%)	2(10%)	20(19,42%)
*S-P2.6	7(58,33%)	5(41,67%)		12(11,65%)

Nota. Elaboración propia.

Campo de problema 3: Reconocer la distribución de probabilidad de una variable aleatoria continua como una herramienta que permite ver la variación aleatoria

Paralelamente al paso dos, es importante trabajar con la función de probabilidad de una variable aleatoria continua, que permite cuantificar la posibilidad de ocurrencia que esta variable tome sus valores en un subconjunto de números reales. Por consecuencia este campo de problema se divide en las siguientes cuatro situaciones-problemas.

S-$P_{3.1}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua. Se exponen actividades que a partir de una función de densidad (lenguaje simbólico) es necesario construir su gráfica, aquello es ejemplificado en la Figura 2.9, pregunta 1.

Figura 2.9

Ejemplo de situación-problema 3.1 y 3.3 (T5)

■ Considera la función $f(x) = 0,2$ definida en el intervalo [0, 5].

1. Construye la gráfica de **f**.
2. Demuestra que **f** puede definir una función de densidad de una variable aleatoria continua.

Nota. Extraída del Departamento de Investigaciones Educativas (2014, p.254).

S-P$_{3.2}$ Calcular algunas probabilidades asociadas a una variable aleatoria continua. Se observan tareas que por medio de la función de densidad (lenguaje simbólico o gráfico) se solicita determinar la probabilidad de que la variable aleatoria continua involucrada tome valores en un intervalo de la recta real (ver Figura 2.10).

Figura 2.10

Ejemplo de situación-problema 3.2 (T5)

1. Una estación de servicio tiene dos bombas, cada una de las cuales puede dispensar hasta 10.000 galones de gasolina por mes. La cantidad de gasolina dispensada en un mes es una variable aleatoria, en miles de galones, y su función de distribución es:

$$f(x) = \begin{cases} x & \text{si } 0 < x < 1 \\ 2 - x & \text{si } 1 \leqslant x \leqslant 2 \\ 0 & \text{en otro punto} \end{cases}$$

Calcula la probabilidad de que las bombas dispensen entre 0,5 y 0,9 miles de galones el siguiente mes.

A. 0,4
B. 0,5
C. 0,9
D. 0,28
E. 0,36

Nota. Extraída del Departamento de Investigaciones Educativas (2014, p.238).

S-P$_{3.3}$ Determinar si la función dada es una función de probabilidad de una variable aleatoria continua. Se proponen actividades que presentan una función definida en un intervalo de la recta real y es necesario corroborar si esta cumple las condiciones para ser función de densidad, como se ejemplifica en la Figura 2.9, pregunta 2.

**S-P$_{3.4}$ Calcular el valor de una incógnita para que la función dada sea de probabilidad*. Se exponen tareas que proponen una función de densidad y se requiere calcular el valor de una incógnita presente en su expresión algebraica, de manera que la función satisfaga las condiciones para ser de densidad. La Figura 2.11, ejemplifica la situación-problema.

Figura 2.11

Ejemplo de situación-problema 3.4 (T5)

7. ¿Para qué valor de **a** la función f(x) = a, definida en el intervalo]2, 5[, podría ser una función de densidad para una variable aleatoria continua?

A. $\frac{1}{2}$
B. $\frac{1}{3}$
C. $\frac{1}{4}$
D. $\frac{1}{5}$
E. $\frac{1}{7}$

Nota. Extraída del Departamento de Investigaciones Educativas (2014, p.241).

La Tabla 2.5 sintetiza el análisis desarrollado a los libros sobre el tercer campo de problema. En ella se expone el predominio de las actividades sobre la S-$P_{3.2}$ calcular probabilidades asociadas a una variable aleatoria continua (41,86%) y S-$P_{3.3}$ determinar si la función propuesta es de densidad (37,21%). Debido a que la definición de función de densidad involucra el concepto de integral, el cual no forma parte del currículo chileno, en los textos escolares la metodología adoptada para trabajar con aquella función es por medio de la interpretación geométrica de la integral. Por tanto, las dos situaciones-problemas señaladas involucran la representación gráfica de la función de densidad. También, fue evidenciada la presencia de las tareas correspondientes a la S-$P_{3.4}$ calcular el valor de una incógnita para que la función propuesta sea de probabilidad (9,3%), la cual no es sugerida en el currículo chileno vigentes.

Tabla 2.5

Frecuencias y porcentajes sobre el tercer campo de problema en textos chilenos

Situaciones-Problemas	10°(n=0)	11°(n=0)	12°(n=17)	Total(n=17)
S-P3.1			5(100%)	5(11,63%)
S-P3.2			18(100%)	18(41,86%)
S-P3.3			16(100%)	16(37,21%)
*S-P3.4			4(100%)	4(9,3%)

Nota. Elaboración propia.

Campo de problema 4: Establecer la función de distribución de una variable aleatoria discreta

En un tercer paso, es necesario estudiar la función de distribución, debido a que es importante para el cálculo de probabilidades y es un tópico inseparable de la variable aleatoria. Actualmente en el currículo escolar chileno solo se trabaja con aquella función para variables aleatorias discretas y para ello fueron identificadas tres situaciones-problemas:

S-$P_{4.1}$ Determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta. Se observan actividades que proponen un experimento aleatorio y variable aleatoria discreta, o una función de probabilidad (lenguaje simbólico, tabular o gráfico), y es solicitado calcular la probabilidad de un subconjunto de valores de la variable involucrada. La Figura 2.12 ejemplifica esta situación-problema.

Figura 2.12

Ejemplo de situación-problema 4.1 (T1)

3. La siguiente función de probabilidad de la variable aleatoria X: "número de mascotas que tiene una familia en cierta ciudad del país", se muestra en la siguiente tabla:

x_i	0	1	2	3
$P(X = x_i)$	0,05	0,45	0,15	0,35

c. ◆ Se escoge al azar una familia. ¿Cuál es la probabilidad de que esta tenga al menos 1 mascota? ¿Y que tenga a lo más 2 mascotas?

Nota. Extraída de Díaz et al. (2020, p.144).

S-P$_{4.2}$ Definir la función de distribución de una variable aleatoria discreta. Se proponen tareas que presentan un experimento aleatorio y variable aleatoria discreta, o una función de probabilidad (lenguaje tabular), y es necesario calcular la probabilidad que la variable involucrada tome, a lo más, cada uno de sus valores posibles (ver Figura 2.13).

Figura 2.13

Ejemplo de situación-problema 4.2 (T3)

Una urna tiene 2 monedas de un peso, 2 de 5 pesos y una de 10 pesos. Si se extraen 3 monedas al azar, y se define la v.a. discreta X: cantidad extraída en pesos.

5. Determina la función de distribución de probabilidad de **X**.

Nota. Extraída de Blanco et al. (2009, p.325).

S-P$_{4.3}$ Representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta. Se presentan tareas en las que, por medio de un experimento aleatorio y variable aleatoria discreta, o una función de probabilidad (lenguaje tabular, gráfico o simbólico), se requiere construir el gráfico de la función de distribución asociado a la variable involucrada, como se ejemplifica en la Figura 2.14.

Figura 2.14

Ejemplo de situación-problema 4.3 (T3)

Se extraen 3 cartas sin remplazo de una baraja de naipes españoles (40 cartas). Se definen las siguientes v.a. discretas:
X: número de ases extraídos.
Y: número de figuras extraídas.

2. Grafica las funciones de probabilidad y distribución de probabilidad de **X** e **Y**. ¿Dónde se concentran los datos?

Nota. Extraída de Blanco et al. (2009, p.325).

La Tabla 2.6 resume la indagación desarrollada sobre los textos que mencionan el cuarto campo de problema. En ella predominan las actividades correspondientes a la S-$P_{4.1}$ cálculo de probabilidad a un subconjunto de valores de una variable aleatoria discreta (58,97%), por sobre la S-$P_{4.3}$ representación de la función de distribución en lenguaje gráfico (21,79%) y S-$P_{4.2}$ representación de dicha función en leguaje simbólico (19,23%). Además, existe coherencia entre las situaciones-problemas identificadas en las orientaciones curriculares chilenas (Bizet et al., 2023) y aquellas observadas en los libros.

Tabla 2.6

Frecuencias y porcentajes sobre el cuarto campo de problema en textos chilenos

Situaciones-Problemas	10°(n=24)	11°(n=11)	12°(n=0)	Total(n=35)
S-P4.1	35(76,09%)	11(23,91%)		46(58,97%)
S-P4.2	11(73,33%)	4(26,67%)		15(19,23%)
S-P4.3	10(58,82%)	7(41,18%)		17(21,79%)

Nota. Elaboración propia.

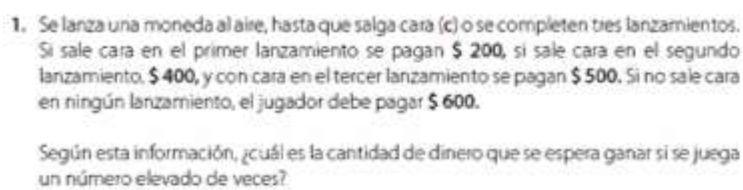
Campo de problema 5: Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria

En un cuarto paso es esencial trabajar con la media, desviación estándar y varianza de una variable aleatoria, las cuales están mayormente relacionadas a su función de probabilidad. Para ello este campo de problema fue descompuesto en las siguientes seis situaciones-problemas.

S-$P_{5.1}$ Calcular la media o el valor esperado de una variable aleatoria discreta. Se exponen actividades que proponen un experimento aleatorio y variable aleatoria discreta o función de probabilidad (lenguaje tabular, simbólico o gráfico), con el propósito de calcular el valor esperado de la variable involucrada (ver Figura 2.15).

Figura 2.15

Ejemplo de situación-problema 5.1 (T4)



1. Se lanza una moneda al aire, hasta que salga cara (c) o se completen tres lanzamientos. Si sale cara en el primer lanzamiento se pagan **$ 200**, si sale cara en el segundo lanzamiento, **$ 400**, y con cara en el tercer lanzamiento se pagan **$ 500**. Si no sale cara en ningún lanzamiento, el jugador debe pagar **$ 600**.

   Según esta información, ¿cuál es la cantidad de dinero que se espera ganar si se juega un número elevado de veces?

Nota. Extraída de Blanco et al. (2009, p.154).

S-$P_{5.2}$ Interpretar la media o el valor esperado de una variable aleatoria discreta. Se observan tareas que proporcionan un experimento aleatorio y variable aleatoria discreta, o una función de probabilidad (lenguaje tabular o gráfico), con el objetivo de calcular e interpretar el valor esperado de la variable involucrada, como se ejemplifica en la Figura 2.16.

Figura 2.16

Ejemplo de situación-problema 5.2 (T4)

2. Un juego consiste en sacar una bolita de una urna que contiene **7** bolitas rojas y **3** azules. Gana **$ 500** si la bolita que se extrae es de color rojo y el jugador debe pagar **$ 1.500** en caso de que la bolita sea azul. ¿Es conveniente jugar?

Nota. Extraída de Blanco et al. (2009, p.155).

S-P$_{5.3}$ *Calcular la varianza de una variable aleatoria discreta*. Se presentan actividades en donde a través de un experimento aleatorio y variable aleatoria discreta, o función de probabilidad (lenguaje tabular, simbólico o gráfico), es necesario calcular la varianza de la variable asociada. Aquello es ejemplificado en la Figura 2.17, pregunta 13.

Figura 2.17

Ejemplo de situación-problema 5.3, 5.4 y 5.5 (T4)

En la siguiente tabla se muestra el pronóstico de la rentabilidad (en millones de pesos) para los próximos 7 meses de los bancos A y B.

X e Y corresponden a la v.a. "rentabilidad en millones de pesos" para los bancos A y B, respectivamente. $P(X = x)$ y $P(Y = y)$ son las funciones de probabilidad para X e Y.

$X = Y$	$P(X = x)$	$P(Y = y)$
44	0,04	0,02
55	0,07	0,003
59	0,03	0,03
62	0,04	0,04
65	0,07	0,087
69	0,72	0,72
71	0,03	0,1

13. Calcula la varianza y desviación estándar para ambos bancos.

14. ¿Qué banco posee la mayor rentabilidad?

Nota. Extraída de Blanco et al. (2009, p.157).

S-P$_{5.4}$ *Calcular la desviación estándar de una variable aleatoria discreta*. Se proponen tareas que, por medio de un experimento aleatorio y variable aleatoria discreta, o función de probabilidad (lenguaje simbólico o tabular), se solicita hallar el valor de la desviación estándar de aquella variable involucrada, la Figura 2.17, pregunta 13, ejemplifica la situación-problema.

S-P$_{5.5}$ *Interpretar la desviación estándar de una variable aleatoria discreta*. Se muestran actividades que entregan una función de probabilidad (lenguaje tabular), con la finalidad de determinar la desviación estándar de esta variable e interpretar su valor, esta es ejemplificada en la Figura 2.17, pregunta 14.

S-$P_{5.6}$ *Identificar la media y desviación estándar de una variable aleatoria continua.* Se presentan tareas que proponen en lenguaje simbólico una distribución normal o una variable aleatoria continua distribuida normalmente, y es requerido reconocer su media y desviación estándar. La Figura 2.18 ejemplifica la presente situación-problema.

Figura 2.18

Ejemplo de situación-problema 5.6

e. Si se tiene una población que se distribuye de forma normal modelada por *N*(2, 1), ¿cuál es la media y la desviación típica?

Nota. Extraído de Osorio et al. (2019, p.174).

La Tabla 2.7 sintetiza el análisis realizado sobre los textos que tratan el quinto campo de problema. Aquella muestra que se fomentaron más las tareas relativas a calcular algunos valores de posición central o de dispersión de una variable aleatoria discreta (S-$P_{5.1}$ y S-$P_{5.3}$), como la media (34,55%) y varianza (21,82%) respectivamente, por sobre la interpretación de estos, dado que existe baja presencia de las actividades sobre la S-$P_{5.2}$ interpretar la esperanza de una variable aleatoria discreta (14,55%) y S-$P_{5.5}$ interpretar su desviación estándar (5,45%). También en este contexto existe coherencia entre las situaciones-problemas reconocidas en el currículo escolar chileno (Bizet et al., 2023) y las observadas en los libros.

Tabla 2.7

Frecuencias y porcentajes sobre el quinto campo de problema en textos chilenos

Situaciones-Problemas	10°(n=0)	11°(n=21)	12°(n=7)	Total(n=28)
S-P5.1		15(78,95%)	4(21,05%)	19(34,55%)
S-P5.2		5(62,5%)	3(37,5%)	8(14,55%)
S-P5.3		9(75%)	3(25%)	12(21,82%)
S-P5.4		7(87,5%)	1(12,5%)	8 (14,55%)
S-P5.5		2(66,67%)	1(33,33%)	3(5,45%)
S-P5.6			5(100%)	5(9,09%)

Nota. Elaboración propia.

Campo de problema 6: La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real

En un quinto paso, es fundamental trabajar con la aplicación de la variable aleatoria en los modelos teóricos de distribuciones de probabilidad, en un primer momento los relativos a variables de tipo discreta, aunque en las actuales orientaciones curriculares chilenas solo es estudiado el modelo binomial, para el cual se distinguieron cinco situaciones-problemas.

S-$P_{6.1}$ *Identificar situaciones que pueden modelizarse a través de una distribución binomial*. Se plantean tareas que presentan un experimento aleatorio dicotómico y variable aleatoria discreta, con el objetivo de reconocer si esta posee distribución binomial, como es ejemplificado en la Figura 2.19, pregunta b.

Figura 2.19

Ejemplo de situación-problema 6.1, 6.2, 6.3 y 6.4 (T2)

3. Analiza el siguiente experimento y responde:

Se lanza doce veces un dado de seis caras y se define la variable aleatoria
X: número de caras obtenidas que son múltiplos de 3.

a. Determina el valor de los parámetros *p*, *q* y *n*.
b. ¿Reúne las condiciones para ser modelado mediante la distribución binomial? Justifica tu respuesta.
c. ¿Cuál es la probabilidad de obtener a lo más 4 éxitos?
d. ¿Cuál es la probabilidad de obtener al menos 5 éxitos? ¿Cómo se relaciona con la probabilidad de obtener a lo más 4 éxitos?
e. ¿Cuál es el valor esperado y la varianza del número de éxitos? Interprétalos.

Nota. Extraído de Osorio et al. (2019, p.169).

S-$P_{6.2}$ *Determinar los parámetros asociados a una distribución binomial*. Se proponen actividades que solicitan hallar el valor de los parámetros n, p y/o q, a partir de distinta información proporciona en su enunciado: (i) un experimento aleatorio y variable aleatoria; (ii) algunas medidas de dispersión; (iii) probabilidad. Esta situación-problema se ejemplifica en la Figura 2.19, pregunta a.

S-$P_{6.3}$ *Calcular la media y varianza de una distribución binomial*. Se observan tareas que, por medio de un experimento aleatorio y variable aleatoria discreta, se requiere calcular la media y/o varianza para la distribución binomial que modela la situación propuesta (ver Figura 2.19, pregunta e).

S-$P_{6.4}$ *Determinar probabilidades asociadas a una distribución binomial de forma manual*. Se exponen tareas que, a partir de un experimento aleatorio y variable aleatoria binomial, es necesario calcular la probabilidad de un valor o un conjunto de valores de aquella variable realizando el trabajo manualmente. La Figura 2.19, preguntas c y d, ejemplifican aquello.

S-P$_{6.5}$ Determinar probabilidades asociadas a una distribución binomial empleando una herramienta tecnológica. Se sugieren actividades que, a través de la tecnología (softwares o calculadoras gráficas) como herramienta de trabajo sea requerido simular un experimento aleatorio y/o calcular probabilidades asociadas a un modelo binomial.

La Tabla 2.8 resume la indagación desarrollada sobre los textos respecto al sexto campo de problema. Esta muestra que predominaron las tareas correspondientes a la S-P$_{6.1}$ identificar situaciones modelizadas a través de una distribución binomial (28,57%) y S-P$_{6.4}$ determinar probabilidades asociadas a este modelo de forma manual (44,44%), principalmente en grado 11 con el 77,78% y 71,43%, respectivamente. También expone que no se fomentó el uso de tecnología como herramienta de trabajo para estudiar la distribución binomial, dado que se presentó inexistencia de la actividad S-P$_{6.5}$ utilizar tecnología en el cálculo de probabilidades.

Tabla 2.8

Frecuencias y porcentajes sobre el sexto campo de problema en texto chilenos

Situaciones-Problemas	10°(n=0)	11°(n=15)	12°(n=9)	Total(n=24)
S-P6.1		14(77,78%)	4(22,22%)	18(28,57%)
S-P6.2		4(40%)	6(60%)	10(15,87%)
S-P6.3		1(14,29%)	6(85,71%)	7(11,11%)
S-P6.4		20(71,43%)	8(28,57%)	28(44,44%)
S-P6.5				

Nota. Elaboración propia.

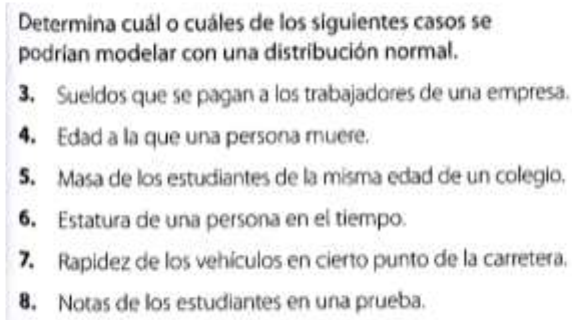
Campo de problema 7: La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real

Paralelamente al quinto paso, es importante estudiar aplicaciones de la variable aleatoria en modelos teóricos de distribuciones de probabilidad de referencias para variables de tipo continuas, puntualmente en el currículo chileno se sugiere el modelo normal, para el cual fueron distinguidas ocho situaciones-problemas.

S-P$_{7.1}$ Identificar situaciones que pueden modelizarse a través de una distribución normal. Se proponen tareas que describen una situación experimental donde es necesario reconocer si está podría describir a una variable aleatoria continua en la que la mayoría de los valores se agrupan en torno a un valor central y los valores extremos son menos frecuentes (ver Figura 2.20).

Figura 2.20

Ejemplo de situación-problema 7.1 (T5)



Determina cuál o cuáles de los siguientes casos se podrían modelar con una distribución normal.

3. Sueldos que se pagan a los trabajadores de una empresa.
4. Edad a la que una persona muere.
5. Masa de los estudiantes de la misma edad de un colegio.
6. Estatura de una persona en el tiempo.
7. Rapidez de los vehículos en cierto punto de la carretera.
8. Notas de los estudiantes en una prueba.

Nota. Extraída del Departamento de Investigaciones Educativas (2014, p.227).

S-P$_{7.2}$ *Determinar probabilidades asociadas a una distribución normal de forma manual.* Se observan actividades que proponen una variable aleatoria continua distribuida normalmente o una situación modelada por medio de la distribución normal (lenguaje verbal o gráfico), con media y desviación típica conocida, y se requiere calcular probabilidades relacionadas con la variable involucrada realizando el trabajo manualmente. Aquello se ejemplifica en la Figura 2.21

Figura 2.21

Ejemplo de situación-problema 7.2 (T2)

6. Los tiempos de espera (en minutos) en un paradero de dos empresas de buses interurbanos son modelados con las siguientes distribuciones normales:

Distribución $N(12, 3)$ Distribución $N(8, 2)$

Según esa información, ¿en cuál existe mayor probabilidad de superar los 5 minutos de espera? Argumenta.

Nota. Extraída de Osorio et al. (2019, p.191).

S-P$_{7.3}$ *Describir la tendencia de los datos representados gráficamente, empleando la distribución normal o la aproximación de la binomial por la normal*. Se presentan actividades en las que por medio de un software se requiere simular un experimento aleatorio y graficar la distribución de probabilidad de una variable aleatoria, con el propósito de describir aquella distribución o que ocurre con esta a medida que aumenta el número de repeticiones del experimento. Esta situación-problema se ejemplifica en la Figura 2.22.

Figura 2.22

Ejemplo de situación-problema 7.3 (T5)

Considera ahora el experimento de lanzar el mismo dado 50 veces y además la variable aleatoria X del ejemplo.

X cantidad de veces que se obtiene 1 punto.

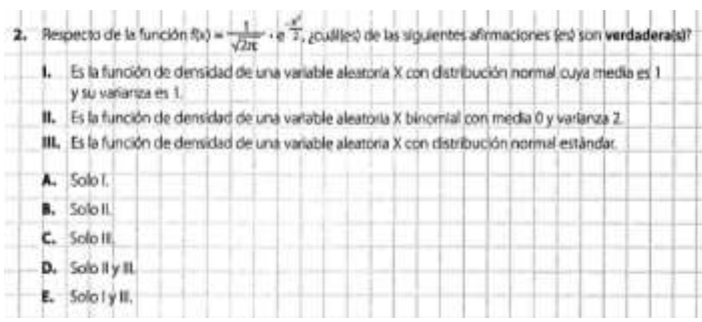
6. Utiliza el *software* para simular la distribución de probabilidades de la variable aleatoria en este caso. Describan la gráfica que consiguieron.

Nota. Extraída de Departamento de Investigaciones Educativas (2014, p. 236).

S-P$_{7.4}$ Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal. Se proponen tareas que entregan afirmaciones sobre el modelo normal con el objetivo de determinar su veracidad o falsedad, la Figura 2.23 ejemplifica aquella situación-problema.

Figura 2.23

Ejemplo de situación-problema 7.4 (T5)



2. Respecto de la función $f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$, ¿cuál(es) de las siguientes afirmaciones (es) son **verdadera(s)**?

I. Es la función de densidad de una variable aleatoria X con distribución normal cuya media es 1 y su varianza es 1.

II. Es la función de densidad de una variable aleatoria X binomial con media 0 y varianza 2.

III. Es la función de densidad de una variable aleatoria X con distribución normal estándar.

A. Solo I.

B. Solo II.

C. Solo III.

D. Solo II y III.

E. Solo I y III.

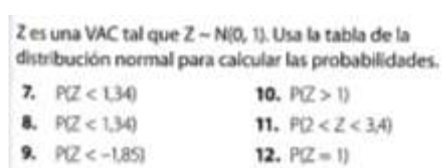
Nota. Extraída de Departamento de Investigaciones Educativas (2014, p.239).

S-P$_{7.5}$ Determinar probabilidades asociadas a una distribución normal empleando una herramienta tecnológica. Se sugieren actividades que, a través del uso de la tecnología (software o calculadoras gráficas), sea solicitado determinar probabilidades vinculadas a un modelo normal e interpretar los resultados proporcionados por la herramienta utilizada.

**S-P$_{7.6}$ Calcular la probabilidad en una distribución normal estándar.* Se plantean tareas que a partir de una variable aleatoria continua distribuida normalmente con media 0 y desviación típica 1, es requerido determinar probabilidades relacionadas a aquella variable (ver Figura 2.24).

Figura 2.24

Ejemplo de situación-problema 7.6 (T5)



Z es una VAC tal que Z ~ N(0, 1). Usa la tabla de la distribución normal para calcular las probabilidades.

7. $P(Z < 1{,}34)$ **10.** $P(Z > 1)$

8. $P(Z < 1{,}34)$ **11.** $P(2 < Z < 3{,}4)$

9. $P(Z < -1{,}85)$ **12.** $P(Z = 1)$

Nota. Extraída del Departamento de Investigaciones Educativas (2014, p.229).

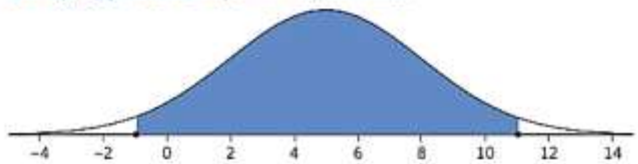
**S-P7.7 Representar en lenguaje gráfico una distribución normal*. Se exponen actividades que por medio de una variable aleatoria continua distribuida normalmente con media y desviación típica concreta, se debe construir la curva normal con las características entregadas, la Figura 2.25 ejemplificada esta situación-problema.

Figura 2.25

Ejemplo de situación-problema 7.7 (T2)

2. Observa el ejemplo. Luego, representa gráficamente y calcula las probabilidades en tu cuaderno.

Si $X \sim N(5, 3)$, para calcular $P(-1 < X < 11)$ tendremos:



Remplazando:

$\mu - 2\sigma = 5 - 2 \cdot 3 = -1$

$\mu + 2\sigma = 5 + 2 \cdot 3 = 11$

Tendremos que $P(-1 < X < 11) = P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) = 0{,}954$

a. Si $X \sim N(14, 3)$, calcula $P(11 < X < 17)$.

b. Si $Y \sim N(45, 7)$, calcula $P(Y < 31)$.

c. Si $Z \sim N(126, 8)$, calcula $P(Z > 150)$.

Nota. Extraída de Osorio et al. (2019, p.175).

**S-P7.8 Calcular los valores correspondientes a una probabilidad dada*. Se observan tareas que presentan una variable aleatoria continua distribuida normalmente o una situación modelada por una distribuir normal, con media y desviación típica específica, con el propósito de obtener el conjunto de valores de la variable (intervalo real) correspondiente a una probabilidad dada, como es ejemplificado en la Figura 2.26.

Figura 2.26

Ejemplo de situación-problema 7.8 (T5)

Considera las variables aleatorias $X \sim N(1, 2)$ y $Z \sim N(0, 1)$.
Encuentra valores de a para que se cumpla lo pedido.

13. $P(Z < a) = 0{,}5$	**16.** $P(X > a) = 0$
14. $P(Z > a) = 0{,}5$	**17.** $P(X > a) = 0{,}5$
15. $P(Z < a) = 0{,}05$	**18.** $P(X > a) = 0{,}05$

Nota. Extraída del Departamento de Investigaciones Educativas (2014, p. 231).

La Tabla 2.9 sintetiza el análisis realizado sobre los textos que tratan el séptimo campo de problema. Esta muestra predominio de las actividades que involucran calcular probabilidades de forma manual, como la S-P7.2 vinculada a una distribución normal (54,28%) y S-P7.6 asociada a una distribución normal estándar (18,57%). Además, destacan las tareas de realizar el proceso inverso al señalado, asociadas a la S-$P_{7.8}$ calcular valores correspondientes a una probabilidad dada (10,71%). Respecto a las dos últimas situaciones-problemas mencionadas (7.6 y 7.8), aquellas no son sugeridas en el currículo chileno vigente, por tanto, existe desarmonía entre este documento y los libros analizados.

Tabla 2.9

Frecuencias y porcentajes sobre el septimo campo de problema en textos chilenos

Situaciones-Problemas	10°(n=0)	11°(n=0)	12°(n=57)	Total(n=57)
S-P7.1			6(100%)	6(4,29%)
S-P7.2			76(100%)	76(54,28%)
S-P7.3			4(100%)	4(2,86%)
S-P7.4			7(100%)	7(5%)
S-P7.5				
*S-P7.6			26(100%)	26(18,57%)
*S-P7.7			6(100%)	6(4,29%)
*S-P7.8			15(100%)	15(10,71%)

Nota. Elaboración propia.

También la Tabla 2.9 expone que no fueron fomentadas las actividades correspondientes a la S-$P_{7.5}$ determinar probabilidades asociadas al modelo normal mediante herramientas tecnológicas. Aunque en el contexto de las tareas sobre la S-$P_{7.3}$ describir la tendencia de los datos representados gráficamente empleando la distribución normal o aproximación de ella, presente en baja proporción (2,86%), fue propuesto utilizar software para simular experimentos aleatorios y graficar el modelo señalado. Esta última acción muestra cierta coherencia con las orientaciones curricular vigentes, donde se promueve el uso de tecnología para estudiar la normal.

Por otra parte, se observó baja presencia de las actividades relacionadas a la S-$P_{7.7}$ representar en lenguaje gráfico el modelo normal (4,29%), lo cual puede deberse a que esta situación-problema no es sugerida en los lineamientos curriculares actuales.

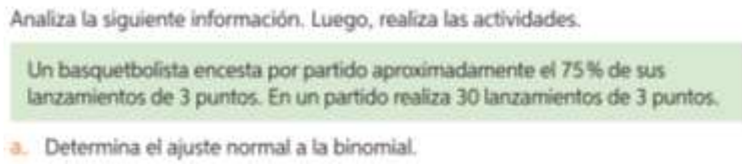
Campo de problema 8: Aproximar distribuciones de variables aleatorias discretas con un número grande de valores

En el último paso, es importante continuar estudiando aplicaciones de la variable aleatoria en distribuciones de probabilidad, específicamente la aproximación de la binomial por la normal, para lo cual este campo de problema fue descompuesto en dos situaciones-problemas.

S-$P_{8.1}$ Calcular los parámetros asociados a una distribución normal. Se plantean tareas que presenta una situación modelada por una distribución binomial donde el experimento dicotómico involucrado fue repetido un número grande de veces, con el objetivo de comprobar que los parámetros n y p satisfacen simultáneamente algunas condiciones y determinar la media y desviación estándar de la distribución normal que se aproxima a la binomial propuesta. La Figura 2.27 ejemplifica esta situación-problema.

Figura 2.27

Ejemplo de situación-problema 8.1 (T2)


Analiza la siguiente información. Luego, realiza las actividades.

Un basquetbolista encesta por partido aproximadamente el 75 % de sus lanzamientos de 3 puntos. En un partido realiza 30 lanzamientos de 3 puntos.

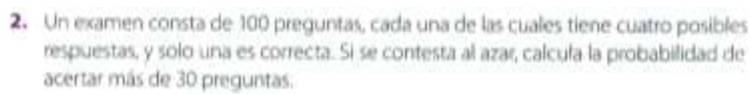
a. Determina el ajuste normal a la binomial.

Nota. Extraída de Osorio et al. (2019, p.186).

S-P$_{8.2}$ Calcular probabilidades en el contexto de aproximación de la distribución binomial a la normal. Se proponen actividades que, por medio de una situación modelada por una distribución binomial, posible de aproximarse a una normal, es requerido calcular probabilidades (ver Figura 2.28).

Figura 2.28

Ejemplo de situación-problema 8.2 (T5)


2. Un examen consta de 100 preguntas, cada una de las cuales tiene cuatro posibles respuestas, y solo una es correcta. Si se contesta al azar, calcula la probabilidad de acertar más de 30 preguntas.

Nota. Extraída del Departamento de Investigaciones Educativas (2014, p.234).

La Tabla 2.10 resume el estudio desarrollado sobre los textos que recogen el octavo campo de problema. Aquella muestra predominio de las tareas vinculas a la S-P$_{8.2}$ calcular probabilidades en el contexto de aproximar la distribución binomial por la normal (89,66%), situación-problema propuesta en los lineamientos curriculares chilenos vigentes (Bizet et al., 2023), por tanto, existe coherencia entre esta normativa y los textos analizados.

Tabla 2.10

Frecuencias y porcentajes sobre el octavo campo de problema en textos chilenos

Situaciones-Problemas	10°(n=0)	11°(n=0)	12°(n=17)	Total(n=17)
S-P8.1			3(100%)	3(10,34%)
S-P8.2			26(100%)	26(89,66%)

Nota. Elaboración propia.

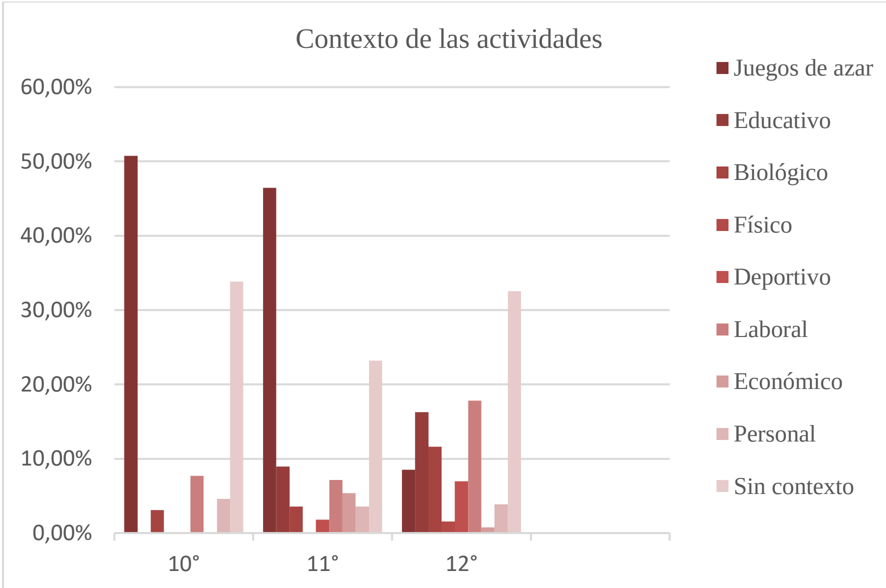
3.2.4.3 Contexto de las actividades

Por último, fue analizado el contexto de cada actividad indagada, donde se encontró gran diversidad e identificó nueve categorías: (a) juego de azar, lanzar dados, monedas o fichas, extraer de una tómbola bolas, girar una ruleta, sacar una carta de un mazo o una pieza de domino del juego; (b) educativo, resolución de una prueba de selección múltiple y aprobación o reprobación de una asignatura; (c) biológico, masa corporal y estatura de personas, duración de los síntomas del resfrío, tipos de sangre de una población, sexo de bebés en una familia, longitud o masa de un animal y eficiencia de un medicamento para este. (d) físico, fenómenos naturales, como la cantidad anual de terremotos en el mundo o cantidad mensual de agua caída en cierta ciudad; (e) deportivo, rendimiento de un deportista, por ejemplo, un tirador con arco en cierta cantidad de disparos, un basquetbolista en un partido, deportistas en una maratón, un equipo de fútbol durante un mes; (f) laboral, productividad de una empresa y falencias cometida en el proceso, como: cantidad de cajas desocupadas en un supermercado o salas en funcionamiento de un cine, tiempo que demora un cliente en ser atendido en una empresa o satisfacción de este, horas que tarda un artesano en crear un producto, vida útil de un computador, cantidad de fallas en teléfonos o zapatos de una fábrica, número de tornillos defectuosos en un proceso de ensamblaje; (g) económico, rentabilidad de bancos o planes en que se desea invertir, y solicitud de dinero en crédito de un banco; (h) personal, cantidad de mascotas de una familia, horas a la semana qua lee una persona, ceros en su cédula de identidad, cigarros que consume al día y llamadas diarias recibida; (i) sin contexto, actividad que carecen de un contexto.

La Figura 2.29, muestra la distribución de los contextos presentes en las actividades analizadas, según nivel educativo. Esta evidencia que en grado 12 fueron observados todos los contextos previamente descritos, nueve categorías, predominando la descontextualizada (32,55%), laboral (17,82%) y educativa (16,28%). También en grado 11 se reconoció gran diversidad de contextos, ocho categorías, destacando los juegos de azar (46,43%) y sin contexto (23,21%). Mientras que en grado 10 se identificaron algunos contextos, cinco categorías, prevaleciendo juego de azar (50,76%) y descontextualizada (33,85%).

Figura 2.29

Contextos utilizados en las actividades presentes en los textos analizados



Nota. Elaboración propia.

3.2.5 Conclusiones

El análisis desarrollado permitió mostrar ¿cómo se aborda la variable y sus aplicaciones sobre distribuciones de probabilidad en libros de texto escolares?, para ello se indagaron actividades presentes en cinco libros de texto dirigido a estudiantes entre los grados 10 y 12, desde la perspectiva del EOS. Así fue observado que hay situaciones-problemas fomentadas en los libros de texto analizados que no fueron sugeridas en el currículo chileno, tales como: diferenciar entre variables aleatorias discretas y continuas, calcular el valor de incógnitas de manera que la función propuesta sea de probabilidad (para ambos tipos de variables señaladas), calcular probabilidades en una distribución normal estándar, representar en un gráfico la distribución normal y calcular los valores correspondientes a una probabilidad dada en el contexto de la distribución normal.

También fue evidenciado que existen situaciones-problemas identificadas en el currículo chileno que no fueron abordadas en los textos escolares o poseen mínima presencia, como: calcular probabilidades asociadas a distribuciones binomial y normal empleando una herramienta tecnológica, y diferenciar entre variables aleatorias y variables con dependencia funcional, respectivamente. Por tanto, los resultados demuestran falta de alineación entre la propuesta del MINEDUC (2015; 2019a) y los textos analizados. Además, muestran un escenario opuesto al sugerido en: investigaciones (Choo-Kim y Choo-Peng, 2015), donde la tecnología es una buena herramienta para mejorar la comprensión de las distribuciones binomial y/o normal en estudiantes; y lineamientos curriculares internacionales (NCTM, 2000), que recomiendan realizar tareas sobre simulación utilizando software para construir distribuciones de probabilidad empíricas y luego comparar estos resultados con las distribuciones de probabilidad teóricas, como por ejemplo la normal y binomial.

La ausencia o escasa presencia de dichas situaciones-problemas, podría conllevar que los estudiantes posean dificultades vinculadas a un inadecuado tratamiento de los tópicos en cuestión, así como también que los profesores interpreten incorrectamente que aquellas tareas no son relevantes de trabajar en educación escolar.

Respecto a la diversidad de situaciones-problemas sobre la variable aleatoria, en los libros analizados fueron observadas 24 tipos, organizadas en cinco campos de problemas (C-P_1 a C-P_5). Por tanto, aquella realidad chilena promueve el estudio de variable aleatoria en la escuela y está en consonancia con directrices curriculares internacionales que recomienda fomentar su conocimiento entre los conceptos probabilísticos básicos (Franklin et al., 2005; NCTM, 2000). Además estos resultados son más alentadores para la educación probabilísticas escolar que los obtenidos por Ortiz (2002), quien propuso nueve situaciones-problemas y observó en libros españoles de bachillerato presencia de cuatro de ellas: determinar probabilidades asociadas a los valores de una variable aleatoria discreta (desde un enfoque clásico), definir su función de probabilidad, construir un gráfico de barras para representa aquella función, y calcular el valor de la variable aleatoria discreta al que corresponde una cierta probabilidad.

Así mismo en los documentos indagados fue evidenciado gran variedad de situaciones-problemas sobre la aplicación de la variable aleatoria en distribuciones de probabilidad. Particularmente para el modelo binomial fueron observadas cuatro tipos, estructurada en un campo de problema (C-P_6), resultado novedoso pues no se han encontrado estudios relativo al tratamiento de aquel tópico en textos escolares. Mientras que para el modelo normal se constaron 9 tipos, organizadas en dos campos de problemas (C-P_7 y C-P_8), lo que demuestra mejores resultados que Valverde (2017), la cual comprobó la existencia de a lo más cuatro situaciones-problemas en libros españoles de bachillerato.

Por otra parte, fue observado una alta variedad de contextos en las actividades analizadas, concretamente nueve categorías, escenario que concuerda con lineamientos curriculares internacionales (NCTM, 2000), los cuales proponen que las tareas en todos los niveles escolares deben incluir trabajar en diversos contextos fuera de las matemáticas. Puntualmente en grado 10, cuando es introducido el estudio de la variable aleatoria, predominan tareas en el ámbito de juego de azar, resultados similares que los observados por Doukhan y Gueudet (2019) en textos franceses**,** donde destacan las con contexto no-matemático.

En tanto que en los grados 11 y 12, cuando son abordados en los textos los modelos binomial y normal, se manifiesta gran diversidad de contextos, ocho y nueve respectivamente, sobresaliendo las categorías de juego de azar, sin contexto y laboral. Aquel resultado en torno a la normal está en consonancia con el obtenido por Valverde (2017) en libros españoles, donde predomina el último contexto señalado entre los cinco que propuso.

Finalmente, estudios como este, en torno a libros de texto, ayudan a mejorar la enseñanza de las matemáticas (Marco-Buzunáriz et al., 2016). Puntualmente se espera que este sea un insumo valioso para los profesores, que deben enseñar en escuelas los tópicos de variable aleatoria y/o distribuciones binomial y normal y construir experiencias de enseñanza, los investigadores interesados en desarrollar estudios sobre su comprensión, y quienes diseñan textos escolares. Debido a que el presente trabajo proporciona criterios para mejorar la elaboración de libros de texto, los cuales continúa siendo el recurso didáctico más empleado para enseñar matemáticas en la educación escolar (Vásquez y Alsina, 2015).

Referencias

Amrani, H. y Zaki, M. (2015). Student's conceptual difficulties with respect to the notion of random variable. *International Journal of Education, Learning and Development, 3*(9), 82-98.

Batanero, C. (2004) Ideas estocásticas fundamentales ¿qué contenido se debe enseñar en la clase de probabilidades? En J. Fernández, M. Sousa y S. Ribeiro (Eds.), *Actas do I Encontro de Probabilidades e Estatística na Escola* (pp. 9-30). Centro de Investigação em Educação da Universidade do Minho.

Bizet, V., Molina-Portillo, E., Ruz, F. y Contreras, J.M. (2023). Elaboración de una guía de situaciones-problemas sobre variable aleatoria y sus aplicaciones a partir del currículo escolar chileno. *Educación Matemática, 35*(1), 163-190.

Blanco, M., Bozt, J., Calderón, F., Jiménez, M., González, M., López, G., Romero, P., Díaz, M., Muñoz, G. y Rupin, P. (2009). *Matemática 2 proyecto bicentenario*. Santillana.

Blanco, M., Bozt, J., Calderón, F., Romero, L., Jiménez, L. y Jammet, C. (2009). *Matemática 3 proyecto bicentenario.* Santillana.

Burrill, G. y Biehler, R. (2011). Fundamental Statistical Ideas in the School Curriculum and in Training Teachers. En C. Batanero, G. Burrill, y C. Reading (Eds.), *Teaching Statistics in School*

Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study (pp.57-69). Springer.

Choo-Kim, T. y Choo-Peng, T. (2015). Effects of the handheld technology instructional approach on performances of students of different achievement levels. *Computers and Education*, *82,* 306-314.

Departamento de Investigaciones Educativas. (2014). *Matemática 4 proyecto bicentenario*. Santillana.

Díaz, E., Ortiz, N., Morales, K., Rebolledo, M., Barrera, R. y Norambuena, P. (2020). *Texto del estudiante de matemática 2° medio*. Ediciones SM.

Dinges, H. (2005). Variables, in particular random variables. En M. Hoffmann, J. Lenhard, y F. Seeger (Eds.*), Activity and sign grounding mathematics education* (pp. 305–311). Springer.

Doukhan, C. y Gueudet, G. (2019). Students' difficulties at the secondary-tertiary transition: the case of random variables. En U. Jankvist, M. van den Heuvel-Panhuizen y M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 2464-2471). Freudenthal Group y Freudenthal Institute of the Utrecht University y ERME.

Font, V. y Godino, J. (2006). La noción de configuración epistémica como herramienta de análisis de textos matemáticos: su uso en la formación de profesores. *Educação Matemática Pesquisa*, *8*(1), 67-98.

Font, V., Godino, J. y Gallardo, J. (2013). The emergence of objects from mathematical practices. *Educational Studies in Mathematics, 82*(1), 97-124.

Franklin, C., Kader, G., Newborn, D., Moreno, J., Peck, R., Perry, M. y Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-k–12 curriculum framework*. American Statistical Association.

García-García, J., Fernández, N., Arredondo, E. e Imilpán, I. (2022). The Binomial Distribution: Historical Origin and Evolution of Its Problem Situations. *Mathematics*, *10*, 1-28.

Gea, M. (2014). *La correlación y regresión en bachillerato: análisis de libros de texto y del conocimiento de los futuros profesores* [Tesis de Doctorado, Universidad de Granada].

Godino, J. (1999). Implicaciones metodológicas de un enfoque semiótico- antropológico para la investigación en didáctica de las matemáticas. En T. Ortega (Ed.*), Investigación en Educación Matemática III* (pp. 196-212). SEIEM.

Godino, J., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education*, *39*(1-2), 127-135.

Godino, J., Batanero, C. y Font, V. (2019). The onto-semiotic approach: implications for the prescriptive character of didactics. *For the Learning of Mathematics, 39*(1), 37- 42.

Heitele, D. (1975). An epistemological view of fundamental stochastic ideas. *Educational Studies in Mathematics*, *6*(2), 187–205, 1975.

Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la investigación*. Editorial McGraw Hill Education.

Krippendorff, K. (1990). *Metodología de análisis de contenido: teoría y práctica*. Paidós.

Marco-Buzunáriz, M. A., Muñoz-Escolano, J. M. y Oller-Marcén, A. M. (2016). Investigación sobre libros de texto en los simposios de la SEIEM (1997-2015). En J. Macías, A. Jiménez, J. González, M. Sánchez, P. Hernández, C. Fernández, F. Ruiz, T. Fernández y A. Berciano (Eds.), *Investigación en Educación Matemática XX* (pp. 325-334). SEIEM.

Ministerio de Educación de Chile. (2009). *Currículum: Objetivos fundamentales y contenidos mínimos obligatorios de la educación básica y media*. Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2015). *Bases Curriculares 7° básico a 2° medio. Gobierno de Chile.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2016). *Matemática, programa de estudio segundo medio.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019a). Chile. *Bases Curriculares 3º y 4º medio.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019b). *Programa de estudio matemática 3° medio para formación general*. Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019c). *Programa de estudio matemática 4° medio para formación general.* Unidad de Currículum y Evaluación.

Ministerio de Educación, Cultura y Deporte (2015). Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato. *Boletín Oficial del Estado*, *1*, 169-546.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standars for Mathematics.* NGACBP y CCSSO.

Occeli, M. y Valeiras, N. (2013). Los libros de texto de ciencias como objeto de investigación: una revisión bibliográfica. *Enseñanza de las Ciencias*, *31*(2), 133- 152.

Ortiz, J. (2002). *La probabilidad en los libros de texto.* Grupo de Investigación en Educación Estadística de la Universidad de Granada.

Osorio, G., Norambuena, P., Romante, M., Gaete, D., Díaz, J., Celedón, J., Morales, K., Ortiz, N., Ramírez, P., Barrera, R. y Hurtado, Y. (2019). *Texto del estudiante de matemática 3° y 4° medio*. Ediciones SM.

Ruiz, B. (2013). *Análisis epistemológico de la variable aleatoria y comprensión de objetos matemáticos relacionados por estudiantes universitarios* [Tesis de Doctorado, Universidad de Granada].

Stahl, S. (2006). *The Evolution of the Normal Distribution. Mathematics Magazine, 79*(2), 96-113.

Tauber, L. (2001). *La construcción del significado de la distribución normal a partir de actividades de análisis de datos* [Tesis de Doctorado, Universidad de Sevilla].

Valverde, M. (2017). *Un estudio de la presentación de la distribución normal en los textos de bachillerato* [Tesis de Maestría, Universidad de Granada].

Vásquez, C. y Alsina, A. (2015). Un modelo para el análisis de objetos matemáticos en libros de texto chilenos: situaciones problemáticas, lenguaje y conceptos sobre probabilidad. *Profesorado, revista de currículum y formación de profesorado, 19*(2), 441-462.

Vilca, M. (2015). *Tipificación de los errores que se presentan al identificar una variable aleatoria de distribución binomial en problemas contextualizados* [Tesis de Maestría, Pontificia Universidad Católica del Perú].

Zapico, M. (2006). Interrogantes acerca de análisis de contenido y del discurso en los textos escolares. En MINEDUC (Ed.), *Primer seminario internacional de textos escolares (pp.149-155)*. Ministerio de Educación de Chile.

3.3 Estudio 3[4]. Objetos matemáticos ligados a la variable aleatoria y sus aplicaciones sobre distribuciones de probabilidad en libros de texto de educación secundaria chilena

Bizet, V., Molina-Portillo, E. y Contreras, J.M. (2023b). Objetos matemáticos ligados a la variable aleatoria y sus distribuciones de probabilidad en libros de texto chilenos. *PNA, revista de investigación en Didáctica de la Matemática*, *17*(2), 201-238.

Resumen

Este artículo analiza el tratamiento de la variable aleatoria y sus aplicaciones sobre distribuciones de probabilidad en libros de texto de educación secundaria de Chile, desde el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos. Así, un modelo es utilizado para el análisis de objetos matemáticos en libros de texto y currículo escolar chileno, mediante el análisis de contenido. Los resultados demuestran diversidad de lenguajes, conceptos, proposiciones, procedimientos y argumentos ligados a los temas, aunque algunos son identificados en el currículo y excluidos en libros o viceversa, demostrándose falta de coherencia entre los documentos.

Palabras clave: variable aleatoria, distribución binomial, distribución normal, libros de texto, objetos matemáticos, educación escolar.

Abstract

This paper analyzes the treatment of the random variable and its applications on probability distributions in school textbooks in Chile, from the Ontosemiotic Approach to Mathematical Knowledge and Instruction. Thus, a model is used for the analysis of mathematical objects in the Chilean textbooks and the Chilean school curriculum by means of content analysis. The results show a diversity of languages, concepts, propositions, procedures and arguments linked to the topics, although some of them are identified in the curriculum and excluded in books or inversely, showing a lack of coherence between the documents.

[4] Para evitar confusión y repetición en la numeración, se ha modificado la relativa a los epígrafes pertenecientes a los estudios que forman parte del compendio, haciendo referencia al capítulo en que se ubica cada estudio. También se ha cambiado la numeración de Tablas y Figuras con respecto a la versión publicada del artículo, haciendo referencia al estudio en que se enmarcan.

Keywords: random variable, binomial distribution, normal distribution, textbooks, mathematical objects, school education.

3.3.1 Introducción

El estudio de la variable aleatoria y su familia de distribuciones es esencial de abordar en la educación escolar, pues se presentan en experiencias cotidianas (Pfannkuch, 2018). Debido a tal escenario los conceptos estocásticos aludidos han sido incorporados en el currículo de diferentes países entre los grados 9 y 12 (14 a 18 años). Esta tendencia es observada en los *Principles and standards for school mathematics* de la *National Council of Teachers of Mathematics* (NCTM, 2000), que proponen trabajar la variable aleatoria y su media, introducir su función de probabilidad a través de los enfoques de probabilidad frecuencial y clásico, y estudiar la distribución binomial mediante estos dos enfoques. También aquella tendencia es evidenciada en los *Common Core State Standars for Mathematics* de la *National Governors Association Center for Best Practices* y el *Council of Chief State School Officers* (NGACBP y CCSSO, 2010), que sugieren en torno a la variable aleatoria abordar lo señalado e introducir la distribución normal mediante los enfoques frecuencial y clásico.

En tanto los *Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II* (GAISE II) (Bargagliotti et al., 2020) que proporcionan una estructura conceptual para la educación estadística de preescolar a grado 12, en tres niveles de desarrollo de habilidades (A, B y C). Particularmente para el nivel C en el contexto del rol de la probabilidad en estadística, proponen entre los conceptos necesarios de estudiar la variable aleatoria y su media, además de la distribución normal mediante los dos enfoques de probabilidad señalados. Chile no es indiferente a aquella realidad, debido a que las actuales Bases Curriculares del Ministerio de Educación (MINEDUC, 2015; 2019a) introducen la variable aleatoria en grado 10 (15 a 16 años) e incluyen trabajar su función de probabilidad desde el enfoque de probabilidad clásico. En grado 11 (16 a 17 años) sugieren abordar su media, desviación estándar y varianza, y para grado 12 (17 a 18 años) proponen el estudio de las distribuciones binomial y normal bajo los enfoques clásico y frecuencial.

De esta manera, resulta interesante cuestionarse cómo en los libros de texto escolar son abordados los temas aludidos, pues para los profesores aquel recurso representa el currículo y manera de entender y practicar la enseñanza (Escolano, 2009). Sin embargo, son limitados los estudios que indagan en estos documentos el tratamiento de la variable aleatoria y su familia de distribuciones.

En esta perspectiva sobre la variable aleatoria, Ortiz (2002) caracterizó su significado institucional desde el análisis a 11 libros de texto españoles para grado 11. Entre sus resultados propone 11 definiciones y propiedades vinculados a esta, observando en pocos libros a lo más dos propiedades o un concepto, y cuatro lenguajes (verbal, simbólico, gráfico y tabular). También, Doukhan y Gueudet (2019) estudiaron su tratamiento en libros franceses, mediante un análisis comparativo entre dos textos de grado 11 y 12, cuyos resultados muestran que ambos documentos emplearon cuatro diferentes lenguajes, especificados previamente.

Respecto a las distribuciones de probabilidad, Li et al. (2021) analizaron comparativamente el tratamiento de la binomial en seis textos de diferentes países (China, Japón, Singapur, Estados Unidos y Reino Unido), para los grados 10 a 12. Los resultados mostraron que: los libros estadounidense y singapurense abordaron un número mayor de temas en torno al modelo (12) y el chino uno menor (5); los estadounidenses fomentaron mayor diversidad de formas en presentar el contenido (tablas y gráficos, y uso de herramientas tecnológicas), y el chino solo usó tecnología en el contexto de cálculos. Por su parte, Valverde (2017) estableció el significado institucional de la normal a partir del estudio a dos libros de texto españoles de grado 11 y 12, y entre sus resultados observó que: gran parte de los conceptos se presentaron formal y confusamente; fueron promovidos los procedimientos de cálculo de probabilidad y tipificación, el lenguaje verbal, simbólico, tabular y gráfico, y cinco argumentos, predominando la justificación de respuesta mediante ejemplos o contraejemplos y gráfica.

Además, Setiawan (2020) indagó el tratamiento de la variable aleatoria y distribuciones binomial y normal en cinco textos indonesios para grado 12. En sus resultados expuso que: todos los libros proporcionan una definición de variable aleatoria, presentándola uno formalmente como función, tres erróneamente como variable y uno con ideas poco claras; fueron identificados a lo sumo seis temas sobre la binomial en dos textos y solo uno promueve el uso de tabla para el cálculo de probabilidades; han sido reconocidos a lo más diecinueve temas sobre la normal en un libro.

Por otra parte, sobre investigaciones centradas en la comprensión a nivel escolar de los conceptos probabilísticos de interés, se encuentra la desarrollada por Bizet y Ramos (2022), que valoraron en estudiantes chilenos de grado 10 (15 a 16 años) la construcción de variable aleatoria y su función de probabilidad mediante un proceso de instrucción, evidenciando respecto al primer concepto que el 64% de los participantes logró identificarlo y representarlo en lenguaje verbal, gráfico o tabular, y en relación con el segundo un 59%. En tanto Sánchez y Landín (2014) analizaron en estudiantes mexicanos de grado 12 (17 a 18 años) la aplicación de la expresión algebraica de la función de probabilidad de la binomial para resolver dos cuestionarios, evidenciando en el segundo que pocos lograron comprenderla y emplearla adecuadamente, y concluyendo que el diagrama de árbol favorece la comprensión de aquella distribución. Mientras que Valdez y Salinas (2019) indagaron en estudiantes mexicanos de grado 12 (17 a 18 años) la resolución de problemas sobre la normal, observando que lograron desarrollar el procedimiento de estandarización, pero lo hicieron mecánicamente, manifestando dificultad para comprender algunas de sus etapas e identificar la relación entre la probabilidad y el área bajo la curva normal.

Por tanto, el presente estudio tiene el objetivo de analizar el tratamiento otorgado a la variable aleatoria y sus aplicaciones sobre distribuciones de probabilidad en cinco libros de texto chilenos destinados a los grados 9 a 12, bajo la perspectiva del Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos (EOS), debido a que Chile no cuenta con información al respecto.

3.3.2 Marco teórico

La investigación se fundamenta en el EOS (Godino et al., 2020), donde la matemática emerge de la resolución de situaciones-problemas para lo cual se realizan determinadas prácticas matemáticas, entendidas como actuaciones o expresiones verbal, gráfica, etc. desarrollada por una persona o compartidas en una institución (Godino, 2017).

En las prácticas matemáticas surgen e interactúan objetos matemáticos, definidos como entidades materiales o inmateriales que apoyan y regulan la actividad matemática (Godino et al., 2020). En este ámbito, han sido diferenciado seis tipos de objetos matemáticos primarios (Font et al., 2013): situaciones-problemas (tareas, actividades o problemas que promueven la actividad matemática); lenguaje (términos matemáticos en sus diversos registros de representación utilizados para enunciar o resolver tareas); conceptos (descripciones o definiciones relacionadas con un objeto matemático, aplicadas en la resolución de actividades); proposiciones (características o propiedades de conceptos, usadas en solucionar tareas); procedimientos (algoritmos, técnicas de cálculos u operaciones desarrollados para responder actividades);

argumentos (enunciados empleado para aprobar o justificar proposiciones y procedimientos, o responder a tareas).

También, en la actividad matemática, las redes que se establecen de la interacción de aquellos objetos primarios son llamadas configuraciones (Godino et al., 2020): cognitiva, propia de una persona o epistémica, específica de una institución. La noción de configuración epistémica es una herramienta útil para analizar textos matemáticos de diferentes épocas (Font y Godino, 2006) y en la actualidad ha sido empleada en diversas investigaciones sobre educación estadística, por ejemplo, en los trabajos de Alvarado y Batanero (2008), Batanero et al. (2015) y Vásquez y Alsina (2017).

De esta manera, en la presente investigación es empleada la herramienta de configuración epistémica descrita para indagar en cinco textos chilenos el lenguaje, conceptos, proposiciones, procedimientos y argumentos que intervienen en el tratamiento de la variable aleatoria y sus aplicaciones sobre distribuciones de probabilidad. Así este trabajo complementa y completa una investigación previa (Bizet et al., 2023a), donde se identificaron en aquellos libros las situaciones-problemas y sus contextos en torno a los temas señalados, y permite establecer la configuración epistémica que define el significado institucional pretendido de variable aleatoria y sus distribuciones de probabilidad en libros de texto de educación secundaria chilena (grado 9 a 12).

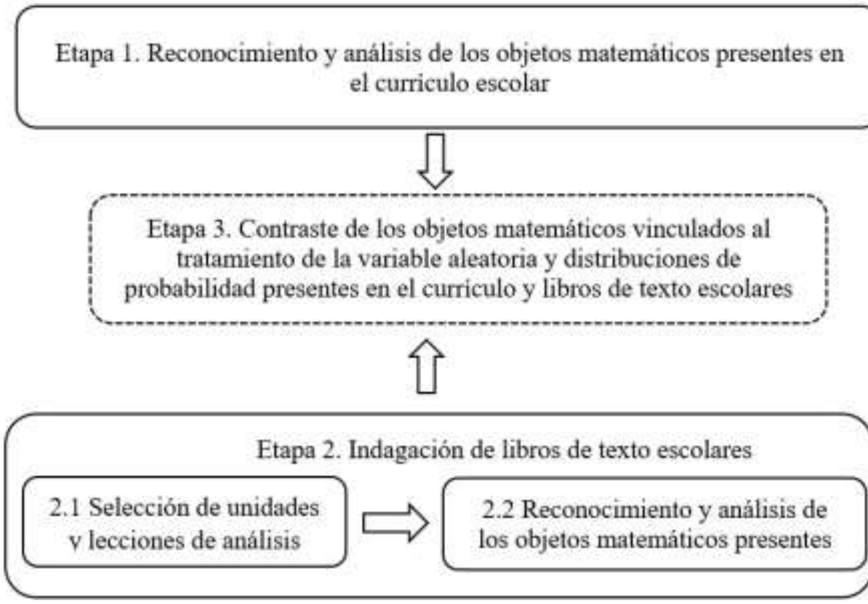
3.3.3 Metodología

Este estudio fue realizado desde una perspectiva cualitativa y es de tipo exploratorio-descriptivo (Hernández et al., 2014). Su desarrollo comprendió tres etapas (ver Figura 3.1) fundamentado en el modelo para el análisis de objetos matemáticos en libros de texto de Vásquez y Alsina (2015).

En la etapa 1 fue empleado como material de investigación el principal documento curricular de Chile propuesto por el Ministerio de Educación, las Bases Curriculares (MINEDUC, 2015; 2019a) para los grados 9 a 12, además del instrumento que hace posible su ejecución denominado Programas de Estudio de Matemática (MINEDUC, 2016; 2019b; 2019c).

Figura 3.1

Etapas desarrolladas en la investigación



Nota. Elaboración propia.

Para la etapa 2, se seleccionó una muestra intencionada, compuesta por dos series de libros de texto chilenos que permanecían vigentes en el curso escolar 2021 (ver Tabla 3.1), donde cada una abarcó los grados 10 (15 a 16 años), 11 (16 a 17 años) y 12 (17 a 18 años). Las series son de reconocidas editoriales del país, la primera consideró los libros utilizados en la educación pública y subvencionada, entregados gratuitamente por el MINEDUC y elaborados por Ediciones SM, donde T1 está en sintonía con las Bases Curriculares dirigidas a grado 10 (MINEDUC, 2015) y T2 responde al escenario vigente de las Bases Curriculares para los grados 11 y 12 (MINEDUC, 2019a). La segunda serie incluyó los textos más usados en la enseñanza privada, los cuales son diseñados por la Editorial Santillana, donde el contenido de variable aleatoria presente en T3 atienden a las Bases Curriculares vigentes para grado 10 (MINEDUC, 2015), y respecto a los contenidos de distribuciones binomial y normal expuesto en T4 y T5, responden al escenario vigente de las Bases Curriculares para los grados 11 y 12 (MINEDUC, 2019a).

Tabla 3.1

Libros de texto utilizados en el análisis

Serie	Código	Nivel educativo	Título	Editorial	Año
1	T1	10°	Texto del estudiante de matemática 2° medio	Ediciones SM	2020
	T2	11° y 12°	Texto del estudiante de matemática 3° y 4° medio		2019

2	T3	10°	Matemática 2 proyecto bicentenario	Santillana	2009
	T4	11°	Matemática 3 proyecto bicentenario		2009
	T5	12°	Matemática 4 proyecto bicentenario		2014

Nota. Elaboración propia.

Tanto en la primera como segunda etapa, los documentos fueron indagados utilizando el análisis de contenido (Krippendorff, 1990) mediante un proceso cíclico e inductivo, abarcando los pasos que se detallan (Cobo, 2003):

1. Distinguir capítulos que tratan los temas de variable aleatoria y/o distribuciones de probabilidad.
2. Indagación de capítulos seleccionados e identificación de los objetos matemáticos primarios (lenguaje: L; conceptos: C; proposiciones: PP; procedimientos: P; argumentos: A), introducidos implícita o explícitamente en el tratamiento de los temas en cuestión.
3. Construcción de tablas que presentan los objetos matemáticos reconocidos en el currículo y textos escolares analizados.

3.3.4 Resultados

3.3.4.1 Etapa 1: Objetos matemáticos presentes en el currículo escolar

El currículo chileno incluyó diversos objetos matemáticos sobre los temas de interés (ver Tabla 3.2), aquellos para grado 10 y 11 principalmente están vinculados a la variable aleatoria (v.a.), mientras que los de último grado a sus aplicaciones en distribución binomial (d.b.) o distribución normal (d.n.).

Este documento para la variable aleatoria promovió cuatro tipos de lenguaje (verbal, simbólico, gráfico y tabular) y cinco métodos de validación (justificación mediante ejemplo o contraejemplo, representación gráfica, simulación con herramienta tecnológica y objetos manipulables, y razonamiento algebraico). Presentó los demás objetos en orden de complejidad creciente, aunque omitió las proposiciones sobre su caracterización mediante la función de densidad y la función de distribución, el procedimiento de cálculo de probabilidades aplicando su función de distribución y los argumentos mediante razonamiento verbal-deductivo y análisis-síntesis. Además, excluyó abordar la integral, entonces los conceptos de función de distribución, media, varianza y desviación estándar sugirió trabajarlos únicamente para el caso de variables aleatorias discretas.

Tabla 3.2

Lenguaje, conceptos, proposiciones, procedimientos y argumentos vinculados al tratamiento de la variable aleatoria y sus aplicaciones sobre distribuciones en el currículo chileno

Objeto matemático	Nivel 10°	Nivel 11°	Nivel 12°
Lenguaje			
L1 Verbal	v.a.	v.a.	d.b. y d.n.
L2 Simbólico	v.a.	v.a.	d.b. y d.n.
L3 Gráfico	v.a.	v.a.	d.b. y d.n.
L4 Tabular	v.a.	v.a.	d.n.
Conceptos			
C1 Variable aleatoria	v.a.	v.a.	d.b. y d.n.
C2 Función de probabilidad	v.a.	v.a.	
C3 Función de densidad			v.a.
C4 Función de distribución	v.a.		
C5 Valores de posición central o de dispersión asociados a una variable aleatoria: C5.1 Media; C5.2 Varianza; C5.3 Desviación estándar		v.a.	
C6 Distribución binomial: C6.1 Función de probabilidad; C6.2 Función de distribución; C6.3 Media; C6.4 Varianza; C6.5 Desviación estándar			d.b.
C7 Distribución normal: C7.1 Función de densidad; C7.2 Media; C7.3 Desviación estándar; C7.4 Moda; C7.5 Mediana; C7.6 Función de densidad de la normal estándar			d.n.
Proposiciones			
PP1 Caracterización de variable aleatoria y variables con dependencia funcional	v.a.		
PP2 Caracterización de variables aleatorias discreta y continua	v.a.		d.b. y d.n.
PP3 Caracterización de la variable aleatoria mediante			
PP3.1 su función de probabilidad	v.a.	v.a.	
PP3.2 su función de densidad			
PP4 Propiedades de la función de probabilidad	v.a.	v.a.	
PP5 Caracterización de la variable aleatoria mediante su función de distribución			
PP6 Propiedades de la función de densidad			v.a.
PP7 Caracterización de la distribución binomial			d.b
PP8 Caracterización de la distribución normal			d.n
PP9 Propiedades de la función de densidad de la normal			d.n
PP10 Propiedades para el cálculo de probabilidades con distribución normal estándar			d.n
PP11 Propiedad de los intervalos centrales ($\pm 3\sigma$)			d.n
PP12 Condición para aproximar una distribución binomial a una normal			d.n
Procedimientos			
P1 Cálculo de la cardinalidad del dominio de una variable aleatoria discreta mediante combinatoria	v.a.		
P2 Partición disjunta del espacio muestral	v.a.	v.a.	
P3 Determinación del dominio de una variable aleatoria discreta o probabilidad asociada a sus valores utilizando diagrama de árbol	v.a.		
P4 Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de probabilidad	v.a.	v.a.	
P5 Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de distribución			

P6 Cálculo de probabilidad de una variable aleatoria continua aplicando su función de densidad			v.a.
P7 Cálculo de valores de posición central y de dispersión asociados a una variable aleatoria discreta empleando sus expresiones algebraicas		v.a.	
P8 Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial			d.b.
P9 Cálculo de probabilidades aplicando la función de distribución de una distribución binomial			d.b.
P10 Cálculo de media, varianza y desviación estándar de una binomial mediante sus expresiones algebraicas			d.b.
P11 Cálculo de probabilidad empleando la propiedad de los intervalos centrales ($\pm 3\sigma$)			d.n.
P12 Tipificación			d.n.
P13 Cálculo de probabilidades utilizando la tabla de distribución normal estándar			d.n.
P14 Corrección por continuidad			
Argumentos			
A1 Justificación por medio de ejemplo o contraejemplo	v.a.	v.a.	
A2 Justificación a través de la representación gráfica	v.a.		d.b. y d.n.
A3 Justificación mediante simulación con			
A3.1 herramienta tecnológica	v.a.		d.b. y d.n.
A3.2 objetos manipulables***	v.a.		d.b.
A4 Justificación mediante la generalización*			d.b. y d.n.
A5 Razonamiento algebraico	v.a.	v.a.	d.b. y d.n.
A6 Razonamiento verbal-deductivo** y ***			
A7 Justificación mediante análisis-síntesis** y ***			

*no aplica para variable aleatoria / **no aplica para distribución binomial / ***no aplica para distribución normal

Nota. Elaboración propia.

En las distribuciones binomial y normal, el currículo fomentó tres tipos de lenguaje (verbal, simbólico y gráfico) y cuatro métodos de justificación (mediante representación gráfica, simulación con herramienta tecnológica y objetos manipulables, generalización y razonamiento algebraico). Además, incluyó solo en el contexto de la normal el lenguaje tabular y para la binomial la validación mediante simulación con objetos manipulables. Respecto a los restantes objetos primarios, primero presentó los relativos a la binomial seguidos por los de la normal, con ascendiente dificultad, donde excluyó el procedimiento de corrección por continuidad, los argumentos mediante ejemplos o contraejemplos y solo para la normal abordó los conceptos de moda y mediana.

3.3.4.2 Etapa 2.1: Unidades y lecciones indagadas en libros de texto

En los cinco libros indagados se seleccionaron las unidades y lecciones que tratan la variable aleatoria y sus aplicaciones en distribuciones de probabilidad (ver Tablas 3.3 y 3.4). Según expone la Tabla 3.3, en grado 11, T2(1) excluye contenidos sobre variable aleatoria, opuesto con lo explicitado por el currículo para este contexto que sugirió trabajar su media, varianza y desviación estándar. La Tabla 3.4 evidencia que para grado 11, T4 propuso estudiar la distribución binomial, lo que difiere con los lineamientos curriculares que introdujeron su estudio en grado 12.

Tabla 3.3

Unidades y lecciones de serie 1 vinculadas al tratamiento de variable aleatoria y su familia de distribuciones

			Serie 1	
Nivel	Texto	Unidad	Lección	Páginas
10°	T1	Probabilidad y Estadística	-Definición de variable aleatoria -Probabilidad de una variable aleatoria -Gráfica de la distribución de una función de probabilidad	136-147 158-159
11°	T2(1)		No se presenta contenido sobre los temas	
12°	T2(2)	Toma de decisiones en situaciones de incerteza	-Valor esperado y varianza de una variable aleatoria -Distribución binomial -Variable aleatoria continua -Distribución normal -Distribución normal estándar -Aproximación normal a la binomial	165-182 185-191

Nota. Elaboración propia.

Tabla 3.4

Unidades y lecciones de serie 2 vinculadas al tratamiento de variable aleatoria y su familia de distribuciones

			Serie 2	
Nivel	Texto	Unidad	Lección	Páginas
10°	T3	Probabilidad	-Variable aleatoria -Función de probabilidad de una variable aleatoria -Función de distribución de probabilidad -Gráficos de función de probabilidad y distribución de probabilidad -Uso de software	320-335
11°	T4	Probabilidad	-Función de probabilidad de una variable aleatoria discreta -Función de distribución de probabilidad de una variable aleatoria discreta -Uso de software -Valor esperado, varianza y desviación estándar -Modelo de distribución binomial	146-165 168-175
12°	T5	Datos y azar	-Variable aleatoria continua -Distribución de probabilidad normal -Uso de software	222-241 254-261

Nota. Elaboración propia.

La Tabla 3.3 muestra que la serie 1 omitió utilizar herramientas tecnológicas durante el tratamiento de los temas de interés. En tanto la serie 2 (ver Tabla 3.4) promovió utilizar softwares en el estudio de la variable aleatoria (T3 y T4) y distribución normal (T5), lo cual está en consonancia con el currículo chileno que en grado 10 y 12 abordó el mecanismo de validación mediante simulación con herramientas tecnológicas.

3.3.4.3 Etapa 2.2: Objetos matemáticos presentes en libros de texto

La metodología señalada permitió indagar libros de texto chilenos y establecer la presencia o ausencia de objetos matemáticos primarios sobre variable aleatoria y su familia de distribuciones identificados previamente en el currículo escolar chileno. Ahora se exponen los resultados de esta fase, donde son descritas las diversas categorías de acuerdo con el lenguaje, conceptos, proposiciones, procedimientos y argumentos asociados a los temas en cuestión, incorporándose un ejemplo en aquellas que sea necesario para su mejor comprensión. Los objetos primarios que se describen a continuación, junto con las situaciones-problemas identificadas previamente (Bizet et al., 2023a) forman la configuración epistémica que define el significado institucional pretendido de variable aleatoria y distribuciones binomial y normal en textos escolares chilenos.

Lenguaje

En el análisis de los libros fueron identificados cuatro tipos de lenguaje utilizados para presentar y trabajar los temas de interés.

L_1 Lenguaje verbal. Palabras y frases que describen tareas, conceptos o sus propiedades, procedimientos y argumentos necesarios para responderlas. Sobre la variable aleatoria se encontraron diversidad de términos: variable aleatoria discreta; espacio muestral; dominio y recorrido; preimagen; valores de la variable aleatoria; evento o suceso aleatorio; función algebraica; variable aleatoria no discreta; tipo de variable aleatoria; variable aleatoria continua; función de probabilidad; gráfica; probabilidad; función de distribución; función de densidad; área achurada o sombreada; intervalo real; curva; promedio o media; valor esperado o esperanza; varianza; variabilidad; homogeneidad; desviación estándar o típica.

También sobre las distribuciones de probabilidad fueron reconocidos diversos términos: (i) distribución binomial; modelo binomial; variable aleatoria binomial; variable de Bernoulli; experimento de Bernoulli; parámetros; probabilidad; modelo de distribución; función de probabilidad; función de distribución; suceso aleatorio; valor esperado o media; varianza; (ii) distribución normal; media; desviación estándar; curva normal o de densidad; área; probabilidad; datos; intervalo real; función de densidad; simetría; mediana; distribución normal estándar; tipificación o normalización; aproximación; ajuste; corrección por continuidad.

L_2 *Lenguaje simbólico.* Notaciones simbólicas y expresiones algebraicas que posibilitan desarrollar operaciones con conceptos y trabajar con mayor grado de complejidad. Las Tablas 3.5 y 3.6 ejemplifica lo observado en libros sobre notaciones y expresiones algebraicas asociada a la variable aleatoria y sus aplicaciones en distribuciones de probabilidad respectivamente.

Tabla 3.5

Ejemplo de L_2 sobre variable aleatoria

Variable aleatoria					
Concepto	Notación	Expresión algebraica	Concepto	Notación	Expresión algebraica
Variable aleatoria	X	$X(\omega) = x_i$	Función de densidad	f(x)	$f(x) = P(a \leq X \leq b)$
Función de probabilidad	f(x)	$f(x_i) = P(X = x_i)$ Propiedades $f(x_1)+..+f(x_i) = 1$ o $\sum f(x_i) = 1$ $0 \leq f(x) \leq 1$	Media	μ E(X)	$\mu = x_1 \cdot p(x_1)+..+x_n \cdot p(x_n)$ $E(X) = \sum_{i=1}^{k} x_i \cdot P(x_i)$
			Varianza	σ^2 Var(X)	$\sigma^2 = (x_1 - \mu)^2 \cdot p(x_1)+.. +(x_n - \mu)^2 \cdot p(x_n)$ $Var(X) = E(X^2) - (E(X))^2$
Función de distribución	F(x)	$F(x_i) = P(X \leq x_i)$	Desviación estándar	s(X)	$s(X) = \sqrt{Var(X)}$

Nota. *Elaboración propia.*

Tabla 3.6

Ejemplo de L_2 sobre distribuciones binomial y normal

Aplicaciones en distribuciones binomial y normal					
Concepto	Notación	Expresión algebraica	Concepto	Notación	Expresión algebraica
v.a. con distribución binomial	$X \rightarrow B(n, p)$ $X \sim B(n, p)$		v.a. con distribución normal estándar	$Z \sim N(0,1)$	$Z = \frac{X - \mu}{\sigma}$
Función de probabilidad de la binomial	f(x)	$f(x) = C_x^n \cdot p^x \cdot (1 - p)^{n-x}$ $P(X = x) = \binom{n}{x} p^x \cdot q^{n-x}$	Función de densidad de la normal estándar	f(x)	$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ Probabilidad en un intervalo: i. $P(Z \leq a)$ con $a \geq 0$ ii. $P(Z \geq a)$ con $a \geq 0$ iii. $P(Z \leq a)$ con $a < 0$ iv. $P(Z \geq a)$ con $a < 0$ v. $P(a \leq Z \leq b) = P(Z \leq b) - P(Z \leq a)$
Función de distribución de la binomial	F(x)	$F(x) = P(X \leq x) = \sum_{i=0}^{x} C_i^n \cdot p^i \cdot q^{n-i}$			
Media	μ	$\mu = n \cdot p$			
Varianza	σ^2	$\sigma^2 = n \cdot p \cdot q$			
v.a. con distribución normal	$X \sim N(\mu, \sigma)$				
	$f(x)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2}$		$B(n,p) \sim N(\mu, \sigma)$	

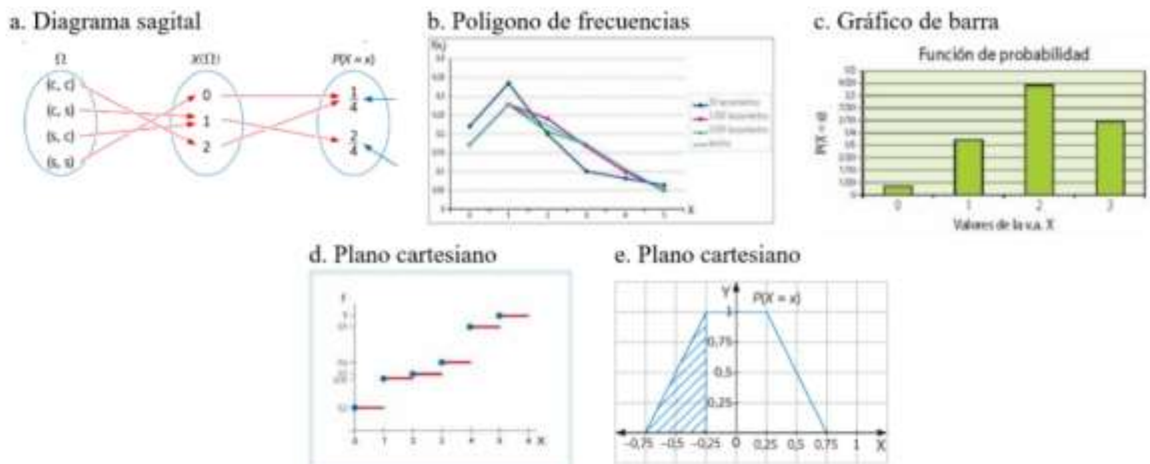
Función de densidad de la normal	Propiedad $\pm 3\sigma$: $P(\mu - \sigma < X < \mu + \sigma) = 0{,}683$ $P(\mu - 2\sigma < X < \mu + 2\sigma) = 0{,}954$ $P(\mu - 3\sigma < X < \mu + 3\sigma) = 0{,}997$	Aproximación binomial por la normal	Corrección por continuidad $P(Y \leq a) = P(X < a + 0{,}5)$ $P(Y < a) = P(X < a - 0{,}5)$ $P(Y \geq a) = P(X > a - 0{,}5)$ $P(Y > a) = P(X > a + 0{,}5)$

Nota. Elaboración propia.

L_3 Lenguaje gráfico. Entre los objetos de la estocástica se reconocen los gráficos. En los libros indagados la variable aleatoria fue representada en un diagrama sagital (Figura 3.2a). Su función de probabilidad se graficó empleando también dicho diagrama (Figura 3.2a), el polígono de frecuencias (Figura 3.2b) y el gráfico de barras (Figura 3.2c); y su función de densidad fue graficada en el plano cartesiano (Figura 3.2e). Su función de distribución se representó en el plano cartesiano (Figura 3.2d).

Figura 3.2

Ejemplo de L_3 sobre variable aleatoria

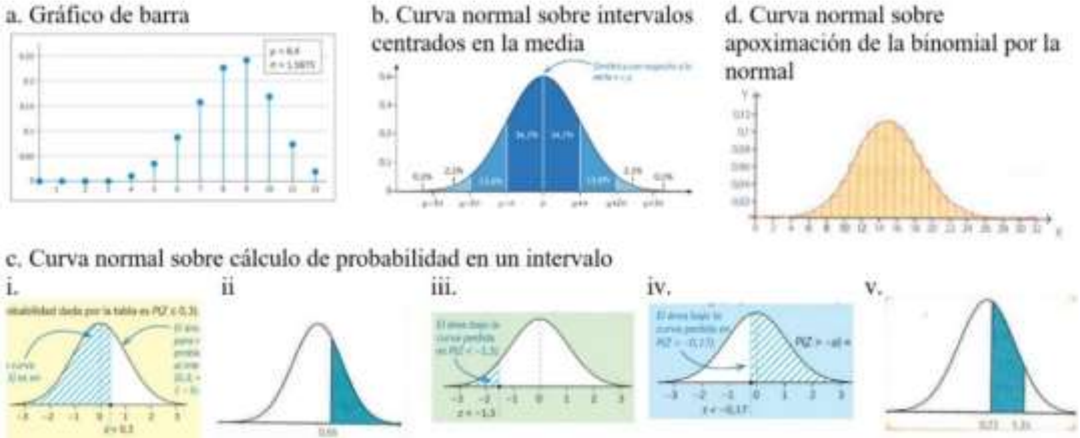


Nota. Elaboración propia.

Respecto al tratamiento de la familia de distribuciones en los textos inspeccionados, la función de probabilidad de la binomial se representó en un gráfico de barras (Figura 3.3a), mientras que la función de densidad de la normal fue graficada con la curva normal, representando distintos contextos: intervalos centrados en la media (Figura 3.3b), cálculo de probabilidades en un intervalo (Figura 3.3c) o aproximación de la binomial por la normal (Figura 3.3d).

Figura 3.3

Ejemplo de L_3 sobre distribuciones binomial y normal

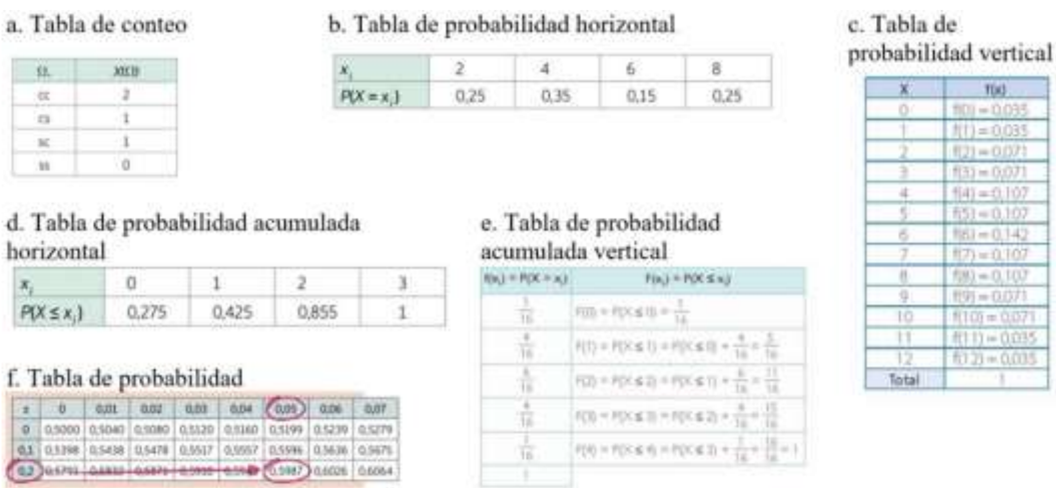


Nota. Elaboración propia.

L_4 Lenguaje tabular. Otro objeto característico de la estocástica son las tablas. En los libros indagados se identificó que la variable aleatoria fue representada en una tabla de conteo (Figura 3.4a). Para representar su función de probabilidad se construyeron tablas de probabilidad horizontal (Figura 3.4b) y vertical (Figura 3.4c), mientras que su función de distribución se representó en una tabla de probabilidad acumulada horizontal (Figura 3.4d) y vertical (Figura 3.4e). En el tratamiento de la función de densidad de la normal estándar esta fue representada en una tabla de probabilidad (Figura 3.4f).

Figura 3.4

Ejemplo de L_4 sobre variable aleatoria y aplicaciones en distribuciones binomial y normal



a. Tabla de conteo

Ω	$X(\omega)$
cc	2
cs	1
sc	1
ss	0

b. Tabla de probabilidad horizontal

x_i	2	4	6	8
$P(X = x_i)$	0,25	0,35	0,15	0,25

c. Tabla de probabilidad vertical

X	f(x)
0	f(0) = 0,035
1	f(1) = 0,035
2	f(2) = 0,071
3	f(3) = 0,071
4	f(4) = 0,107
5	f(5) = 0,107
6	f(6) = 0,142
7	f(7) = 0,107
8	f(8) = 0,107
9	f(9) = 0,071
10	f(10) = 0,071
11	f(11) = 0,035
12	f(12) = 0,035
Total	1

d. Tabla de probabilidad acumulada horizontal

x_i	0	1	2	3
$P(X \leq x_i)$	0,275	0,425	0,855	1

e. Tabla de probabilidad acumulada vertical

$f(x_i) = P(X = x_i)$	$F(x_i) = P(X \leq x_i)$
$\frac{1}{16}$	$F(0) = P(X \leq 0) = \frac{1}{16}$
$\frac{4}{16}$	$F(1) = P(X \leq 1) = P(X \leq 0) + \frac{4}{16} = \frac{5}{16}$
$\frac{6}{16}$	$F(2) = P(X \leq 2) = P(X \leq 1) + \frac{6}{16} = \frac{11}{16}$
$\frac{4}{16}$	$F(3) = P(X \leq 3) = P(X \leq 2) + \frac{4}{16} = \frac{15}{16}$
$\frac{1}{16}$	$F(4) = P(X \leq 4) = P(X \leq 3) + \frac{1}{16} = \frac{16}{16} = 1$
1	

f. Tabla de probabilidad

z	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07
0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,3636	0,5675
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064

Nota. Elaboración propia.

La Tabla 3.7 expone las representaciones sobre variable aleatoria y su familia de distribuciones (binomial y normal) identificadas en los libros indagados, las cuales demostraron una complejidad semiótica de los temas a nivel escolar. Sobre la variable aleatoria se apreció que tanto la serie 1 como 2, promovieron cuatro lenguajes para su representación. En relación con la binomial, T2(2) fomentó mayor variedad de representaciones que T4, aunque ninguno empleó una tabla de valores de su función de probabilidad, que podría facilitar el procedimiento de calcular probabilidades en este contexto. Por último, para el tratamiento de la normal, se evidenció que T2(2) y T5 trabajaron los cuatro tipos de lenguaje.

Tabla 3.7

Lenguajes sobre variable aleatoria y distribuciones de probabilidad

Lenguaje	Variable aleatoria						Distribuciones binomial y normal					
	Serie 1			Serie 2			Serie 1			Serie 2		
	T1	T2(1)	T2(2)	T3	T4	T5	T1	T2(1)	T2(2)	T3	T4	T5
L1 Verbal	v.a.		v.a.	v.a.	v.a.	v.a.			d.b. y d.n.		d.b.	d.n.
L2 Simbólico	v.a.		v.a.	v.a.	v.a.	v.a.			d.b. y d.n.		d.b.	d.n.
L3 Gráfico	v.a.		v.a.	v.a.	v.a.	v.a.			d.b. y d.n.			d.n.
L4 Tabular	v.a.		v.a.	v.a.	v.a.	v.a.			d.n.			d.n.

Nota. Elaboración propia.

Conceptos

En los libros se identificó la presencia o ausencia de conceptos involucrados en el tratamiento de la variable aleatoria y su familia de distribuciones, además de la forma en que fueron introducidos. Según Sfard (1991) un concepto matemático es posible definirlo de manera operacional (descripción de fórmula u operación) o estructural (descripción de condiciones y propiedades). Por tanto, se observó si estos objetos fueron abordados mediante ejemplos, operacionalmente, estructuralmente o combinando aquellas formas, reconociéndose un número total de dieciséis conceptos.

C_1 Variable aleatoria. Fue abordada explícitamente en cinco textos, definiéndola correctamente como una función cuyo dominio es el espacio muestral y recorrido los números reales o un subconjunto de ellos. También se realizó la distinción entre las de tipo discreta y continua, afirmando por ejemplo sobre la primera que "pueden asumir una cantidad finita o infinita numerable de valores" (T5, p.222), y respecto a la segunda "su recorrido es un intervalo de números reales" (T1, p.137).

C_2 Función de probabilidad. También llamada función de cuantía o distribución de probabilidad, fue presentada explícitamente en T1, T3 y T4, describiéndola adecuadamente como una función que asocia a cada valor de la variable aleatoria discreta una posibilidad de ocurrencia.

C_3 Función de densidad. Nombre que recibe la función de probabilidad de una variable aleatoria continua, solo T2(2) y T5 la introdujeron como aquella que permite asignar probabilidad al intervalo $[a, b]$ perteneciente al recorrido de la variable aleatoria continua, mediante el cálculo del área bajo la curva de la función entre los valores a y b.

C_4 Función de distribución. Los textos T1, T3 y T4 estudiaron esta para el caso de variables aleatorias discretas, donde fue descrita como la función que proporciona la probabilidad acumulada hasta un valor determinado de la variable, y solo T3 expuso su domino (números racionales) y recorrido (intervalo [0,1]).

C_5 Valores de posición central y de dispersión asociados a una variable aleatoria. El primero corresponden a la media y el segundo a la varianza y desviación estándar, estos simplifican el estudio y comparación de distribuciones de probabilidad, aunque los libros analizados solo trabajaron con aquellos para variables aleatorias discretas.

$C_{5.1}$ Media. También denominado esperanza o valor esperado, fue tratado por T2(2) y T4 mediante su expresión algebraica (ver Tabla 5); además T4 lo definió en lenguaje verbal como "el promedio ponderado de los valores de X con sus respectivas probabilidades" (p.170) y describió una característica para el contexto de juego (p.155): "(i) este es justo si E(X)=0; (ii) injusto si $E(X) < 0$; (iii) favorable sí $E(X) > 0$".

$C_{5.2}$ Varianza. Fue abordada por T2(2) y T4 mediante su expresión algebraica (ver Tabla 5); también T4 la describió en lenguaje verbal como "el promedio de las distancias al cuadrado de los valores en X hasta la media" (p.170), y describió su característica, que indica el grado de dispersión que posee la variable aleatoria respecto a su valor esperado.

$C_{5.3}$ Desviación estándar. Trabajada por T2(2) en lenguaje verbal como "la raíz cuadrada de la varianza" (p.165) y T4 en lenguaje simbólico mediante su expresión algebraica (ver Tabla 5).

C_6 *Distribución binomial.* Este modelo de probabilidad para variables aleatorias discretas fue introducido en T2(2) y T4, mediante las características de un experimento de Bernoulli: (i) genera solo dos resultados; (ii) la probabilidad de obtener uno de ello es *p* y el otro es *1-p*. Luego se señaló que al realizar aquel experimento *n* veces se genera una distribución binomial, presentándose su función de probabilidad ($C_{6.1}$) en lenguaje simbólico (ver Tabla 6). Solo T4 abordó explícitamente su función de distribución ($C_{6.2}$) mediante su expresión algebraica (ver Tabla 6), y lo mismo ocurrió en T2(2) con su media ($C_{6.3}$) y varianza ($C_{6.4}$). Además, en este ámbito ambos textos descartaron trabajar el concepto de desviación estándar ($C_{6.5}$).

C_7 *Distribución normal*. Aquel modelo de probabilidad para variables aleatorias continuas fue tratado en T2(2) y T5 presentando su función de densidad ($C_{7.1}$) por medio de una expresión algebraica (ver Tabla 6), aunque soló el primero explicitó el dominio (números reales) y únicamente el segundo describió su media ($C_{7.2}$) y desviación estándar ($C_{7.3}$). En este aspecto ambos textos excluyeron trabajar los conceptos de moda ($C_{7.4}$) y mediana ($C_{7.5}$).

La distribución normal estándar o tipificada fue abordada por dos textos: en T2(2) implícitamente cuando trabajó el proceso de tipificación de la variable, mientras que T5 presentó explícitamente su función de densidad ($C_{7.6}$) en lenguaje verbal, afirmando que es un caso particular de la normal donde su media es 0 y desviación estándar 1, y lenguaje simbólico mediante una expresión algebraica (ver Tabla 6).

La Tabla 3.8 sintetiza la manera que los libros introdujeron los conceptos descritos. Esta evidencia que el concepto más definido fue variable aleatoria, seguido por sus funciones de probabilidad y distribución. El primero fue presentado en las dos series de forma operacional, estructural y mediante ejemplo, aunque se observó que en grado 12, T2(2) y T5 emplearon solo estas dos últimas. Las funciones mencionadas fueron introducidas por ambas series de tres formas diferentes, pero la primera abordó los contenidos solo en grado 10 (T1), mientras que la segunda entregó mayor continuidad presentándolos en los grados 10 (T3) y 11 (T4).

Tabla 3.8

Forma de introducción de conceptos sobre variable aleatoria y distribuciones de probabilidad

Conceptos	Serie 1			Serie 2		
	T1	T2(1)	T2(2)	T3	T4	T5
C.1 Variable aleatoria	ESO		ES	ESO	SEO	SE
C2 Función de probabilidad	EOS			EOS	SOE	

C3 Función de densidad		OSE			OSE
C4 Función de distribución	EOS		EOS	EOS	
C5 Valores de posición central y de dispersión asociados a una variable aleatoria					
C5.1 Media		O		EOS	
C5.2 Varianza		O		ESO	
C5.3 Desviación estándar		E		ESO	
C6 Distribución binomial					
C6.1 Función de probabilidad		SO		EO	
C6.2 Función de distribución				EO	
C6.3 Media		O			
C6.4 Varianza		O			
C6.5 Desviación estándar					
C7 Distribución normal					
C7.1 Función de densidad		SO			SO
C7.2 Media					E
C7.3 Desviación estándar					E
C7.4 Moda					
C7.5 Mediana					
C7.6 Función de densidad de la normal estándar					SO

E: definición mediante ejemplo / O: definición operacional / S: definición estructural

Nota. Elaboración propia.

También la Tabla 3.8 muestra diferencias entre las dos series de texto cuando introdujeron los conceptos menos definidos, como algunos valores de posición central y de dispersión asociados a una variable aleatoria, donde la primera serie empleó para la media y varianza la forma operacional y para la desviación estándar un ejemplo, mientras que la segunda utilizó las tres formas que propusimos. Respecto al concepto de distribución binomial, la serie 1 presentó su función de probabilidad de forma estructural y operacional, y su media y varianza operacionalmente, y en tanto la serie 2, introdujo sus funciones de probabilidad y distribución de manera operacional y a través de ejemplos. En cuanto a la distribución normal, ambas series presentaron su función de densidad de forma estructural y operacional, aunque solo la segunda introdujo la función de densidad de la normal estándar del mismo modo señalado, y la media y desviación estándar mediante ejemplos.

Proposiciones

Trece características y propiedades sobre variable aleatoria y su familia de distribuciones fueron reconocidas en los libros examinados.

PP₁ Caracterización de la variable aleatoria y variables con dependencia funcional. Explica las característica de cada una: la primera surge de un experimento aleatorio y corresponde a una función cuyo domino es el espacio muestral de aquel experimento; las segundas se originan de un fenómeno determinista y corresponden a una variable independiente “x” y una variable dependiente “y” que se relacionan mediante una función y=f(x), donde a cada valor real de la variable independiente x (perteneciente al dominio de la función) se asocia un único valor real de la variable dependiente y. Solamente se abordó implícitamente en T1.

PP_2 Caracterización de variables aleatorias discreta y continua. Señala la diferencia entre sus recorridos, pues la primera toma un número finito $(x_1, x_2, \dots, x_n)$ o infinito numerable de valores $(x_1, x_2, \dots)$, mientras que la segunda toma todos los valores comprendidos en un intervalo de recta real (un número infinito no numerable de valores). Esta fue propuesta en T1, T2(2), T3 y T5.

PP_3 Caracterización de la variable aleatoria mediante su función de probabilidad ($PP_{3.1}$) o su función de densidad ($PP_{3.2}$). Cada variable aleatoria hereda una medida de probabilidad que la caracteriza y permite asignar probabilidad a sus valores (función de probabilidad) o a sus intervalos (función de densidad), lo que únicamente se omitió en T2(1).

PP_4 Propiedades de la función de probabilidad. Presenta las condiciones que debe satisfacer esta función: (i) $p_j \geq 0$, para todo j; y (ii) $\sum_{j=1}^{k} p_j = 1$ o $\sum_{j=1}^{+\infty} p_j = 1$. Fueron estudiadas en T1, T3 y T4.

PP_5 Caracterización de la variable aleatoria mediante su función de distribución. Cada variable aleatoria define una función que describe la probabilidad de que ella tome valores menores o iguales que uno de sus valores real. Esto se trabajó en T1, T3 y T4.

PP_6 Propiedades de la función de densidad. Expone las condiciones que debe satisfacer esta función: (i) su probabilidad no es negativa; y (ii) el área limitada por su gráfica en su dominio y el eje de la abscisa es 1. Estas propiedades se trabajaron en T2(2) y T5.

PP_7 Caracterización de la distribución binomial. Establece las características del contexto donde el modelo es aplicable: (i) situación que presenta un suceso dicotómico observado *n* veces; y (ii) los posibles resultados en cada ocasión son independientes. Estas se identificaron en T2(2) y solo (i) en T4.

PP_8 Caracterización de la distribución normal. Constituye las características del contexto donde esta distribución es aplicable*:* (i) situación modelada por una variable aleatoria continua; y

(ii) sus valores se agrupan en torno a un valor central y sus valores extremos no son frecuentes. Estas características solo fueron mencionadas en T2(2) y T5.

PP₉ Propiedades de la función de densidad de la normal: (i) Esta es simétrica con respecto a la media; (ii) la media es, además, la mediana y la moda de la distribución; (iii) $x = \mu$, es el punto sobre el eje horizontal donde la curva tiene su máximo; (iv) la curva tiene sus puntos de inflexión en $x = \mu \pm \sigma$; y (v) el eje de abscisas es una asíntota de la función. Dos libros presentaron algunas de estas propiedades: T2(2) a (i), (iv) y parcialmente (ii), y T5 a (i), (v) y parte de (ii).

PP_{10} Propiedades para el cálculo de probabilidades con distribución normal estándar: (i) $P(Z \leq a)$ o $P(Z < a)$ con $a \geq 0$; (ii) $P(Z \geq a)$ o $P(Z > a)$ con $a \geq 0$; (iii) $P(Z \leq a)$ o $P(Z < a)$ con $a < 0$; (iv) $P(Z \geq a)$ o $P(Z > a)$ con $a < 0$; y (v) $P(a \leq Z \leq b) = P(Z \leq b) - P(Z \leq a)$. Todas fueron trabajadas en T2(2) y T5.

PP_{11} Propiedad de los intervalos centrales ($\pm 3\sigma$). También denominada distribución de casos en relación con la desviación típica y media, corresponde a: (i) $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 68{,}3\%$; (ii) $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 95{,}4\%$; (iii) $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 99{,}7\%$. Fue abordada explícitamente en T2(2) e implícitamente en T5.

PP_{12} Condición para aproximar una distribución binomial a una normal. Dada una variable aleatoria $X \sim B(n, p)$, si *n* es un valor grande de manera que cumple: (i) $n \geq 30$; (ii) $n \cdot p > 5$; y (iii) $n \cdot (1 - p) > 5$, entonces $B(n, p) \sim N(n \cdot p, \sqrt{n \cdot p \cdot (1 - p)})$. Únicamente T5 y T2(2) señalaron explícitamente las condiciones (ii) y (iii), y sobre (i) solo el primero mencionó que *n* debe ser un valor grande.

La Tabla 3.9 resume las proposiciones reconocidas en los libros analizados. Esta evidencia que la más abordada fue la caracterización de la variable aleatoria mediante su función de probabilidad, seguido por las características de las variables aleatorias de tipo discreta y continua, ambas incluidas en las dos series. La menos propuesta ha sido la caracterización de variable aleatoria y variables con dependencia funcional. En tanto las proposiciones que aparecen con baja frecuencia fueron las vinculadas a las distribuciones binomial y normal y las propiedades de la función de densidad, presentes únicamente en un documento de cada serie.

Tabla 3.9

Proposiciones sobre variable aleatoria y distribuciones de probabilidad binomial y normal

Proposiciones	Serie 1			Serie 2		
	T1	T2(1)	T2(2)	T3	T4	T5
PP1 Caracterización de variable aleatoria y variables con dependencia funcional	v.a.					
PP2 Caracterización de variables aleatorias discreta y continua	v.a.		v.a. d.b. d.n.	v.a.	d.b.	v.a. d.n.
PP3 Caracterización de la variable aleatoria mediante						
PP3.1 su función de probabilidad	v.a.		v.a.	v.a.	v.a.	v.a.
PP3.2 su función de densidad			v.a.			v.a.
PP4 Propiedades de la función de probabilidad	v.a.			v.a.	v.a.	
PP5 Caracterización de la variable aleatoria mediante su función de distribución	v.a.			v.a.	v.a.	
PP6 Propiedades de la función de densidad			v.a. d.n.			v.a. d.n.
PP7 Caracterización de la distribución binomial			d.b.		d.b.	
PP8 Caracterización de la distribución normal			d.n			d.n
PP9 Propiedades de la función de densidad de la normal			d.n			d.n
PP10 Propiedades para el cálculo de probabilidades con distribución normal estándar			d.n.			d.n
PP11 Propiedad de los intervalos centrales ($\pm 3\sigma$)			d.n			d.n
PP12 Condición para aproximar una distribución binomial a una normal			d.n			d.n

***Nota*.** Elaboración propia.

Procedimientos

En los textos se identificaron catorce algoritmos, técnicas de cálculo u operaciones empleadas para resolver problemas sobre variable aleatoria y su familia de distribuciones.

P_1 Cálculo de la cardinalidad del dominio de una variable aleatoria discreta mediante combinatoria. Aquella técnica de conteo es empleada en determinar la cantidad de elementos del espacio muestral para luego establecer sus elementos o probabilidad asociada a un valor de la variable. Este procedimiento se desarrolló en T1 y T3.

P_2 Partición disjunta del espacio muestral. Realizada para determinar lo sucesos aleatorios elementales asociados a cada valor de una variable aleatoria discreta, lo que fue abordado implícitamente en T1, T3 y T4.

P_3 Determinación del dominio de una variable aleatoria discreta o probabilidad asociada a sus valores utilizando diagrama de árbol. Solamente T1 sugirió elaborar un diagrama de árbol para identificar el espacio muestral vinculado al experimento aleatorio propuesto.

P_4 Cálculo de probabilidades de una variable aleatoria discreta aplicando su función de probabilidad. Permite determinar la posibilidad de ocurrencia de cada elemento de la variable y compararlas. Esto fue promovido en la mayoría de los libros a excepción de T2(1).

P_5 Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de distribución. Proporciona la probabilidad acumulada hasta algún valor de la variable aleatoria, lo cual se trabajó en T1, T3 y T4.

P_6 Cálculo de probabilidad de una variable aleatoria continua aplicando su función de densidad. Calcular probabilidades en este contexto involucra emplear integrales, concepto excluido en el currículo chileno. Así se observó que T2(2) y T5 sugirieron determinar el área bajo la curva de una función de densidad mediante la interpretación geométrica de la integral definida, es decir, cálculo de áreas de figuras geométricas simples.

P_7 Cálculo de valores de posición central y de dispersión asociados a una variable aleatoria discreta empleando sus expresiones algebraicas. Únicamente T4 incluyó determinar la media, varianza y desviación estándar de ese tipo de variable mediante la expresión algebraica correspondiente. En tanto, T2(2) abordó parcialmente este proceso, dado que excluyó el cálculo de la desviación.

P_8 Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial. Solo T2(2) y T4 promovieron utilizar la expresión algebraica de aquella función para determinar probabilidades, omitiendo el uso de una tabla con sus valores.

P_9 Cálculo de probabilidades aplicando la función de distribución de una distribución binomial. Solamente fue incluido explícitamente en T4 mediante el reemplazo de valores en la expresión algebraica de dicha función, mientras que T2(2) lo abordó implícitamente.

P_{10} Cálculo de media, varianza y desviación estándar de una binomial mediante sus expresiones algebraicas. Únicamente T2(2) promovió este procedimiento a través del reemplazo de valores, sobre la situación binomial dada, en la expresión algebraica del parámetro correspondiente (media y/o varianza).

P_{11} Cálculo de probabilidad empleando la propiedad de los intervalos centrales ($\pm 3\sigma$). Corresponde a determinar la probabilidad de que una variable aleatoria distribuida normalmente

pertenezca a un intervalo centrado en su media y radio k =1, 2 y 3 desviaciones estándar, desarrollado solo en T2(2) y T5.

P_{12} Tipificación. La probabilidad asociada a cualquier distribución normal es posible expresarla en términos de la normal estándar. El procedimiento, denominado tipificación o normalización, permite determinar ese valor y consiste en transformar una variable aleatoria $X \sim N(\mu, \sigma)$ en $Z \sim N(0,1)$ mediante una expresión algebraica. Esto solamente fue trabajado en T2(2) y T5.

P_{13} Cálculo de probabilidades utilizando la tabla de distribución normal estándar. Consiste en utilizar (leer) una tabla de valores de la normal estándar para determinar la probabilidad de que una variable aleatoria continua tome valores en un intervalo dado, cálculo que solo fue incluido en T2(2) y T5.

P_{14} Corrección por continuidad. Utilizar la distribución normal como aproximación de la binomial involucra este procedimiento que consiste en cambiar un número discreto a un intervalo real ubicado a 0,5 por arriba o 0,5 por debajo de aquel número. Esto se debe a que se emplea un modelo de probabilidad de variable aleatoria discreta para aproximarlo a uno de variable aleatoria continua. Únicamente fue abordado explícitamente en T2(2) e implícitamente en T5.

La Tabla 3.10 sintetiza los procedimientos identificados en los textos indagados. Entre los más incluidos destacan calcular probabilidades de una variable aleatoria discreta utilizando sus funciones de probabilidad y distribución, y realizar una partición disjunta del espacio muestral. Mientras que los procedimientos observados con baja frecuencia han sido los relativos a las distribuciones binomial y normal, así como también algunos vinculados a la variable aleatoria discreta (establecer sus valores de posición central y de dispersión mediante expresión algebraica y cardinalidad de su dominio empleando combinatoria) y continua (calcular probabilidades empleando su función de densidad). Además, el menos promovido fue determinar el dominio de una variable aleatoria discreta o sus probabilidades utilizando el diagrama de árbol.

Tabla 3.10

Procedimientos sobre variable aleatoria y distribuciones de probabilidad binomial y normal

Procedimientos	Serie 1			Serie 2		
	T1	T2(1)	T2(2)	T3	T4	T5
P1 Cálculo de la cardinalidad del dominio de una variable aleatoria discreta mediante combinatoria	v.a.			v.a.		
P2 Partición disjunta del espacio muestral	v.a.			v.a.	v.a.	
P3 Determinación del dominio de una variable aleatoria discreta o probabilidad asociada a sus valores utilizando diagrama de árbol	v.a.					
P4 Cálculo de probabilidades de una variable aleatoria discreta aplicando su función de probabilidad	v.a.		v.a.	v.a.	v.a.	v.a.
P5 Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de distribución	v.a.			v.a.	v.a.	
P6 Cálculo de probabilidad de una variable aleatoria continua aplicando su función de densidad			v.a.			v.a.
P7 Cálculo de valores de posición central y de dispersión asociados a una variable aleatoria discreta empleando sus expresiones algebraicas			v.a.		v.a.	
P8 Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial			d.b.		d.b.	
P9 Cálculo de probabilidades aplicando la función de distribución de una distribución binomial			d.b.		d.b.	
P10 Cálculo de media, varianza y desviación estándar de una binomial mediante sus expresiones algebraicas			d.b.			
P11 Cálculo de probabilidad empleando la propiedad de los intervalos centrales ($\pm 3\sigma$)			d.n.			d.n.
P12 Tipificación			d.n.			d.n.
P13 Cálculo de probabilidades utilizando la tabla de distribución normal estándar			d.n			d.n.
P14 Corrección por continuidad			d.n			d.n

Nota. Elaboración propia.

Argumentos

En los textos examinados fueron reconocidos seis métodos de validación sobre variable aleatoria, tres en distribución binomial y cinco para distribución normal, que permitieron relacionar los objetos matemáticos anteriores y las situaciones-problemas identificadas por Bizet et al. (2023a).

A_1 Justificación mediante ejemplo o contraejemplo. Emplea un ejemplo con datos específicos para comprobar una proposición o explicar un procedimiento. Este fue abordado por cinco libros en el ámbito de la variable aleatoria, y por tres (T2(2), T4 y T5) en el contexto de distribuciones binomial o normal, ejemplificándose en la Figura 3.5.

Figura 3.5

Ejemplo de A_1 (T5)

En una variable aleatoria continua X, la probabilidad de que esta tome exactamente un valor **a** es 0, es decir, P(X = a) = 0.

Esto se puede interpretar de la siguiente manera. Si se escoge aleatoriamente a una persona en la calle y mides su masa, se podría preguntar cuál es la probabilidad de que la masa de ella sea 76,43 kg. Esta es una situación compleja, ya que al considerar ese número estamos buscando personas cuya masa, en kilogramos, sea exactamente 76,430000... y seguro que al utilizar instrumentos de mayor precisión observemos valores diferentes a los de una balanza común, por ejemplo, 76,431 kg o 76,43001 kg.

Tiene más sentido preguntar cuál es la probabilidad de encontrar a una persona cuya masa se ubique, por ejemplo, entre 75 kg y 80 kg. Es decir, en una variable aleatoria continua lo natural es calcular la probabilidad de que la variable aleatoria se encuentre en un intervalo dado. Intuitivamente se pude notar que a medida que el intervalo se hace cada vez más pequeño, es más difícil encontrar a una persona cuya masa cumpla con ese requisito.

Nota. Extraída del Departamento de Investigaciones Educativas (2014, p.224).

A_2 Justificación a través de la representación gráfica. Usa las características visuales de un gráfico para aprobar o refutar afirmaciones, o validar una proposición o elabora un gráfico para solucionar un problema. Ella fue incluida en cinco textos en el contexto de variable aleatoria (ejemplo en la Figura 3.6) y solo en T2(2) para el ámbito de distribuciones binomial y normal.

Figura 3.6

Ejemplo de A_2 (T4)

En la siguiente figura se muestra el gráfico de la función de probabilidad de una v.a. **X.** Respecto de la función de probabilidad del gráfico, es verdadero que:

I. Rec X = {0, 1, 2, 3, 4}
II. la probabilidad de obtener **1** es igual a la probabilidad de obtener **3.**
III. la probabilidad de obtener a lo más **3** es de **68,75%.**

A. Solo III
B. I y II
C. I y III
D. II y III
E. I, II y III

Nota. Extraída de Blanco et al. (2009, p.153).

A_3 Justificación mediante simulaciones con herramienta tecnológica u objeto manipulable. Consiste en el análisis de resultados obtenidos en simulación de experimentos aleatorios por medio de software ($A_{3.1}$) u objetos como chinche, moneda, dado, aparato de Galton, etc. ($A_{3.2}$), para extraer conclusiones y validar la solución a problemas. Solo el primer método de validación fue promovido por T3 y T4 en el ámbito de la variable aleatoria y por T5 en el contexto de la normal, como muestra la Figura 3.7.

Figura 3.7

Ejemplo de A_3 (T3)

Mediante una planilla de cálculo se simulará el lanzamiento de **2** dados, para calcular la función de probabilidad de la v.a. **X** definida como **X**: valor absoluto de la diferencia de sus caras, en **50, 1.000** y **5.000** lanzamientos.

Calcula el error para cada una de las simulaciones.

¿Qué observas respecto al error de las probabilidades a medida que aumentan los lanzamientos?

En el siguiente gráfico y tabla, se muestran los valores de la función de probabilidad de la v.a. **X**, para **50, 1.000** y **5.000** lanzamientos, las cuales se comparan con la función de probabilidad teórica de la v.a. **X**, indicando la diferencia o error, con respecto al valor teórico.

x	0	1	2	3	4	5
50	0,22	0,34	0,20	0,10	0,08	0,06
1.000	0,17	0,28	0,24	0,17	0,10	0,05
5.000	0,17	0,28	0,21	0,19	0,11	0,05
Teórico	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Error 50	0,05	0,06	[illegible]	[illegible]	[illegible]	[illegible]
Error 1.000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Error 5.000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

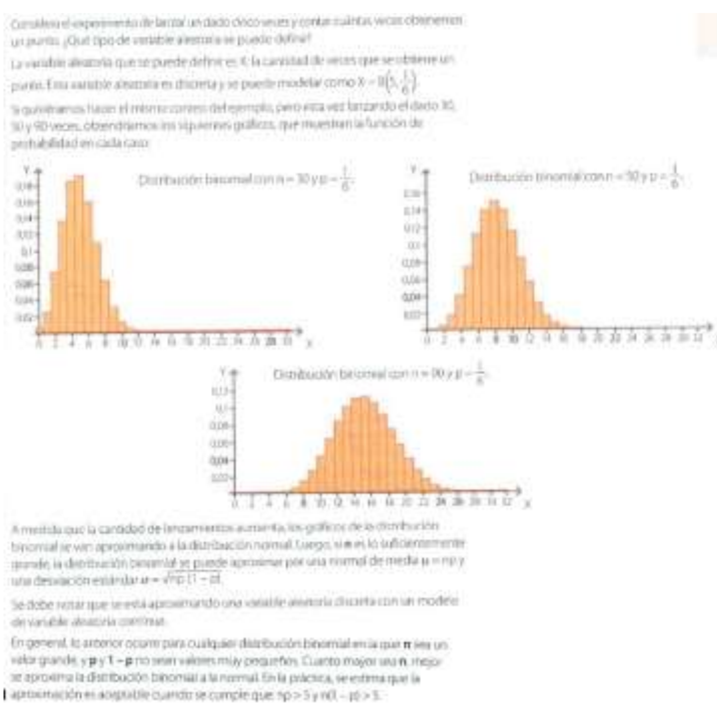
En el gráfico se observa que a medida que aumenta el número de lanzamiento de los dados la probabilidad de cada uno de los valores de la v.a. **X** tienden a ser los mismos. Lo anterior se observa al comparar la diferencia entre la probabilidad experimental y teórica de **X**. En la tabla se observa que el error tiende a disminuir a medida que la cantidad de ensayos aumenta. Este principio se conoce como **ley de los grandes números**, y afirma que al repetir un experimento aleatorio un número muy grande de veces la frecuencia relativa de cada suceso elemental se aproxima a la probabilidad del suceso.

Nota. Extraída de Blanco et al. (2009, p.326-327).

A_4 Justificación mediante la generalización. A partir de ciertas condiciones es desarrollada una demostración informal o formal para obtener conclusiones extensibles a problemas que cumplen con las mismas características iniciales. Solamente fue utilizado para justificar proposiciones por T5 en el ámbito de la normal (ver Figura 3.8).

Figura 3.8

Ejemplo de A_4 (T5)



Considera el experimento de lanzar un dado cinco veces y contar cuántas veces obtenemos un punto. ¿Qué tipo de variable aleatoria se puede definir?

La variable aleatoria que se puede definir es X: la cantidad de veces que se obtiene un punto. Esta variable aleatoria es discreta y se puede modelar como $X \sim B\left(5, \frac{1}{6}\right)$.

Si quisiéramos hacer el mismo [illegible] del ejemplo, pero esta vez lanzando el dado 30, 50 y 90 veces, obtendríamos los siguientes gráficos, que muestran la función de probabilidad en cada caso:

A medida que la cantidad de lanzamientos aumenta, los gráficos de la distribución binomial se van aproximando a la distribución normal. Luego, si **n** es lo suficientemente grande, la distribución binomial se puede aproximar por una normal de media $\mu = np$ y una desviación estándar $\sigma = \sqrt{np(1-p)}$.

Se debe notar que se está aproximando una variable aleatoria discreta con un modelo de variable aleatoria continua.

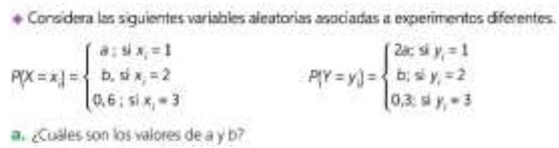
En general, lo anterior ocurre para cualquier distribución binomial en la que **n** sea un valor grande, y **p** y **1 − p** no sean valores muy pequeños. Cuanto mayor sea **n**, mejor se aproxima la distribución binomial a la normal. En la práctica, se estima que la aproximación es aceptable cuando se cumple que: $np > 5$ y $n(1-p) > 5$.

Nota. Extraída del Departamento de Investigaciones Educativas (2014 p.233-234).

A_5 *Razonamiento algebraico.* Involucra la representación, generalización y formalización de regularidades en diversos ámbitos de la matemática, como en probabilidad, con el propósito de ayudar a expresar el pensamiento algebraico (variables, ecuaciones y funciones). Este argumento fue fomentado para justificar procedimientos por cinco libros en el contexto de la variable aleatoria y en tres (T2(2), T4 y T5) para las distribuciones binomial o normal, como se ejemplifica en la Figura 3.9.

Figura 3.9

Ejemplo de A_5 (T1)



◆ Considera las siguientes variables aleatorias asociadas a experimentos diferentes.

$$P(X = x_i) = \begin{cases} a\text{; si } x_i = 1 \\ b,\text{ si } x_i = 2 \\ 0{,}6\text{; si } x_i = 3 \end{cases} \qquad P(Y = y_i) = \begin{cases} 2a\text{; si } y_i = 1 \\ b\text{; si } y_i = 2 \\ 0{,}3\text{; si } y_i = 3 \end{cases}$$

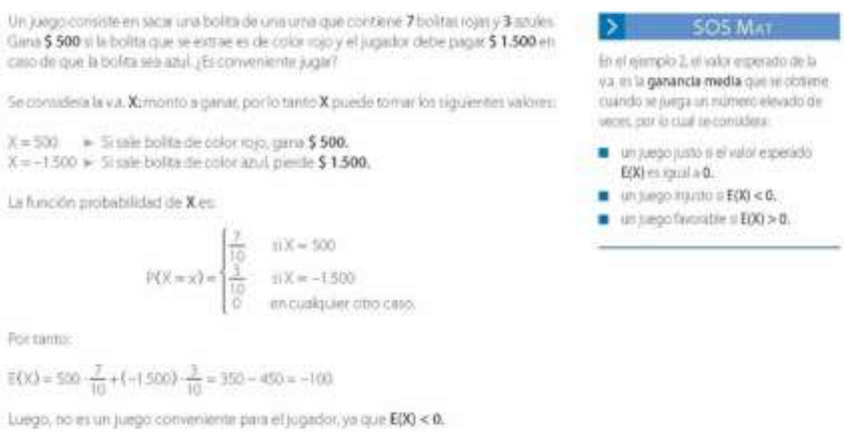
a. ¿Cuáles son los valores de a y b?

Nota. Extraída de Díaz et al. (2020, p.142).

A_6 *Razonamiento verbal-deductivo*. Reconoce un vínculo entre proposiciones que permiten lograr un resultado necesario de interpretar para obtener conclusiones. Fue utilizado por T1 y T4 en torno a la variable aleatoria para justificar procedimientos y respuesta a problemas, como ejemplifica la Figura 3.10.

Figura 3.10

Ejemplo de A_6 (T4)



Un juego consiste en sacar una bolita de una urna que contiene **7** bolitas rojas y **3** azules. Gana **$ 500** si la bolita que se extrae es de color rojo y el jugador debe pagar **$ 1.500** en caso de que la bolita sea azul. ¿Es conveniente jugar?

Se considera la v.a. **X**: monto a ganar, por lo tanto **X** puede tomar los siguientes valores:

X = 500 ► Si sale bolita de color rojo, gana **$ 500.**
X = −1.500 ► Si sale bolita de color azul, pierde **$ 1.500.**

La función probabilidad de **X** es:

$$P(X = x) = \begin{cases} \frac{7}{10} & \text{si } X = 500 \\ \frac{3}{10} & \text{si } X = -1.500 \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

Por tanto:

$$E(X) = 500 \cdot \frac{7}{10} + (-1.500) \cdot \frac{3}{10} = 350 - 450 = -100$$

Luego, no es un juego conveniente para el jugador, ya que **E(X) < 0.**

SOS MAT

En el ejemplo 2, el valor esperado de la v.a. es la **ganancia media** que se obtiene cuando se juega un número elevado de veces, por lo cual se considera:

- un juego justo si el valor esperado **E(X)** es igual a **0.**
- un juego injusto si **E(X) < 0.**
- un juego favorable si **E(X) > 0.**

Nota. Extraída de Blanco et al. (2009, p.155).

A_7 *Justificación mediante análisis-síntesis*. Indaga las cualidades de un problema propuesto, las cuales conducen a concluir propiedades o características de un concepto. Esto fue promovido por T2(2), T3 y T4 en el ámbito de la variable aleatoria y es ejemplificado en la Figura 3.11.

Figura 3.11

Ejemplo de A_7 (T2(2))

Nota. Extraída de Osorio et al. (2019, p.172).

La Tabla 3.11 resume los argumentos reconocidos en los libros examinados. Aquella demuestra que los más promovidos fueron la validación mediante ejemplo o contraejemplo y el razonamiento algebraico, presente en ambas series. Del mismo modo, destaca con alta frecuencia la justificación por medio de representación gráfica. Mientras que los mecanismos de validación menos incluidos han sido la simulación con software y la generalización, solamente observados en la serie 2. Además, ningún texto promovió la justificación a través de simulación con objeto manipulable en el contexto de la variable aleatoria o distribución binomial.

Tabla 3.11

Argumentos sobre variable aleatoria y distribuciones de probabilidad binomial y normal

Argumentos	Variable aleatoria						Distribuciones binomial y normal					
	Serie 1			Serie 2			Serie 1			Serie 2		
	T1	T2(1)	T2(2)	T3	T4	T5	T1	T2(1)	T2(2)	T3	T4	T5
A_1 Justificación por medio de ejemplo o contraejemplo	v.a.		v.a.	v.a.	v.a.	v.a.			d.b. d.n.		d.b.	d.n.
A_2 Justificación a través de la representación gráfica	v.a.		v.a.	v.a.	v.a.	v.a.			d.b. d.n.			
A_3 Justificación mediante simulación con												
$A_{3.1}$ herramienta tecnológica				v.a.	v.a.							d.n.
$A_{3.2}$ objeto manipulable***												
A_4 Justificación mediante generalización*												d.n.
A_5 Razonamiento algebraico	v.a.		v.a.	v.a.	v.a.	v.a.			d.b. d.n.		d.b.	d.n.

A_6 Razonamiento verbal-deductivo** y ***	v.a.			v.a.
A_7 Justificación mediante análisis-síntesis** y ***		v.a.	v.a.	v.a.

*no aplica para variable aleatoria / **no aplica para distribución binomial / ***no aplica para distribución normal

Nota. Elaboración propia.

3.3.4.4 Fase 3: Contraste entre currículo escolar chileno y libros de texto

El análisis desarrollado al currículo y textos escolares sobre los objetos matemáticos demostró que entre los documentos existe coherencia en el lenguaje empleado, dado que para la variable aleatoria y distribución normal propusieron cuatro tipos (verbal, simbólico, gráfico y tabular), y en la binomial tres (verbal, simbólico y gráfico), aunque la segunda serie de libros excluyó el lenguaje gráfico.

Sobre las proposiciones en el contexto de la variable aleatoria y distribución normal se observaron discrepancias entre los documentos analizados. Específicamente en la variable aleatoria el currículo sugirió cinco, excluyendo su caracterización mediante la función de densidad y a través de la función de distribución, mientras que la serie 1 propuso siete, y la serie 2 presentó seis omitiendo la caracterización de variable aleatoria y variables con dependencia funcional. Para la normal el currículo propuso seis, descartando las propiedades de la función de densidad, a diferencia de cada serie que sugirió siete. En tanto que en la binomial aquella normativa sugirió dos al igual que la propuesta por cada serie

Respecto a los conceptos, en el contexto de la variable aleatoria se observó coherencia entre el currículo y las dos series de libros, pues en cada documento se presentaron siete conceptos. En el ámbito de los modelos de probabilidad se observaron discrepancias entre los documentos indagados. Para la distribución binomial el currículo promovió seis conceptos, a diferencia de la primera serie que estableció cuatro (omitiendo su función de distribución y desviación estándar), y la segunda serie tres (omitiendo su media, varianza y desviación estándar). En la normal el currículo sugirió siete, en cambio la serie 1 solo dos, descartando la función de densidad de la normal estándar, media, desviación, moda y mediana, y la serie 2 únicamente presentó cinco, omitiendo las dos últimas señaladas.

También en los procedimientos existió desarmonía entre los documentos analizados. Para la variable aleatoria el currículo promovió seis procedimientos, ignorando calcular probabilidades con su función de distribución. Mientras la serie 1 propuso siete, y la serie 2 presentó seis excluyendo determinar probabilidades o su dominio utilizando diagrama de árbol. En la binomial, el currículo presentó tres procedimientos al igual que la primera serie, pero la segunda serie promovió dos, descartando calcular su media, varianza y desviación estándar mediante expresiones algebraicas. Para el ámbito de la normal el currículo sugirió tres procedimientos exceptuando la corrección por continuidad, mientras que cada serie de libros cuatro.

Por último, en los argumentos se encontraron incongruencias entre los documentos indagados. En la variable aleatoria el currículo fomentó cinco tipos, suprimiendo el razonamiento verbal-deductivo y la justificación mediante análisis-síntesis, a diferencia de los libros: la serie 1 promovió cinco, descartando las simulaciones utilizando herramienta tecnológica y objeto manipulable; y la serie 2 promovió seis, omitiendo el último nombrado. Para la binomial, el currículo presentó cinco argumentos, excluyendo la validación mediante ejemplo o contraejemplo, en tanto la primera serie contempló tres, exceptuando la argumentación a través de generalización y simulaciones con herramienta tecnológica y objeto manipulable, y la segunda serie sólo presentó dos, descartando aquellas tres y la justificación mediante gráfica. En la normal, el currículo promovió cuatro, omitiendo el argumento mediante ejemplo o contraejemplo, mientras que la serie 1 sugirió tres, descartando la validación mediante simulación con herramienta tecnológica y generalización, y la serie 2 contempló cuatro, suprimiendo la justificación por medio de representación gráfica.

3.3.5 Discusión y conclusiones

El análisis realizado permitió mostrar cómo es abordada la variable aleatoria y sus aplicaciones sobre distribuciones de probabilidad en libros de texto escolares chilenos, por ello se analizaron las lecciones sobre estos temas en cinco libros de texto, dirigidos a los grados 10 a 12, desde la perspectiva del EOS. Así se evidenció que existen objetos matemáticos identificados en el currículo chileno que fueron excluidos de los libros de texto, y viceversa, demostrándose una falta de coherencia entre las directrices propuestas por el MINEDUC (2015; 2019a) y aquel recurso. En términos del EOS, la indagación del lenguaje, conceptos, proposiciones, procedimientos y argumentos sobre los temas de interés en los libros chilenos posibilitó el análisis de los principales elementos que contribuyen a la resolución de las situaciones-problemas propuestas en estos textos escolares (Bizet et al., 2023a), donde aquellos objetos matemáticos se organizan formando la configuración epistémica que define el significado institucional pretendido de variable aleatoria y distribuciones binomial y normal en los textos escolares chilenos.

El estudio desarrollado muestra sobre el tratamiento de la variable aleatoria en textos escolares que en el contexto escolar chileno el lenguaje sugerido concuerda con los cuatro tipos observados por Doukhan y Guedet (2019) en el ámbito francés. Este escenario podría favorecer a mejorar el conocimiento de los estudiantes chilenos sobre el concepto en cuestión, es decir, los resultados constatados por Bizet y Ramos (2022), quienes observaron en la resolución a un problema que el 64% de aquellos participantes identificó y representó en lenguaje verbal, figural o tabular la variable aleatoria y el 59% su distribución de probabilidad.

Además, la variedad de conceptos en torno a la variable aleatoria (siete, desde C_1 a $C_{5.3}$) y sus proposiciones (siete, desde PP_1 a PP_6) identificadas en los textos chilenos, evidencia resultados más alentadores que los observados por Ortiz (2002) en el contexto español, quien constató en tales libros solo uno y tres de estos tipos de objetos respectivamente. En Chile, aquella realidad promueve el estudio de variable aleatoria en la escuela y está en consonancia con directrices curriculares internacionales que recomienda fomentar su conocimiento entre los conceptos probabilísticos básicos (Bargagliotti et al., 2020; NCTM, 2000). Sin embargo, el procedimiento de emplear diagrama de árbol fue poco promovido en los libros chilenos, siendo algo que favorece posteriormente la comprensión de la binomial (Sánchez y Landín, 2014).

Por otra parte, este trabajo actualiza el estudio de Li et al. (2021), quienes reconocieron sobre la binomial 12 temas en textos escolares, pues nuestra investigación indagó de manera diferenciada el lenguaje, conceptos, proposiciones, procedimientos y argumentos en torno a aquel modelo. Aunque en este contexto constatamos que los libros chilenos excluyen el lenguaje tabular. La ausencia de la representación tabular de un concepto probabilístico, en el inicio o transición de su estudio, para Lugo-Armenta y Pino-Fan (2021) es una situación preocupante, debido a que no se diversifica los tipos de lenguaje y transiciones entre estos, y limita el desarrollo de habilidades de razonamiento estadístico, siendo este último aspecto una necesidad en todo egresado de educación escolar para afrontar los requerimientos de la ciudadanía (Bargagliotti et al., 2020). Este escenario es contario al resultado de Setiawan (2020) que observó en el ámbito indonesio la utilización del lenguaje tabular de la binomial para calcular probabilidades. Por tanto, se recomienda incluir la tabla de su función de probabilidad y aquel procedimiento que involucra su uso.

Respecto al tratamiento de la normal en libros escolares se concluyen aspectos positivos y negativos. Del lado positivo, fue actualizado el estudio de Setiawam (2020), quien identificó sobre ella 19 temas, pues en este trabajo se diferenció sus objetos matemáticos, identificando en los textos chilenos los mismos cuatro lenguajes observados por Valverde (2017). Un procedimiento más que esta (la corrección por continuidad). Como negativo, se reconoce que los libros analizados promovieron solo introducir conceptos de forma estructural y operacional, similar a lo expuesto en otro trabajo desarrollado en España que señala alto nivel de formalidad (Valverde, 2017), este resultado podría implicar una agudización de aquellos problemas de los estudiantes señalados por Valdez y Salinas (2019) como identificar la relación entre la probabilidad y el área bajo la curva normal y comprender algunas etapas del proceso de estandarización.

También el tratamiento de la normal en los textos indagados incluyó menor variedad de mecanismos de validación que los observados por Valverde (2017), destacando entre los más presentados la validación mediante ejemplos y contraejemplos, concordando con ella, y como menos fomentados la simulación con herramienta tecnológica. Sin embargo, lineamientos curriculares internacionales recomiendan que los estudiantes realicen simulaciones mediante software para construir distribuciones de probabilidad empíricas, comparando estos resultados con las distribuciones de probabilidad teóricas (NCTM, 2000), como por ejemplo la normal y binomial, debido a que la tecnología es una buena herramienta para mejorar la comprensión de estas dos distribuciones (Choo-Kim y Choo-Peng, 2015).

Este estudio completa trabajos previos referidos a las situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad presentes en: i) el currículo escolar chileno (Bizet et al., en 2023b) que en conjunto representan el significado institucional de referencia de los tópicos de interés (ver Anexo 1); ii) textos escolares chilenos (Bizet et al., 2023a) que unidos representan el significado institucional pretendido de los temas en cuestión (Ver Anexo 1). Por tanto, la información presentada es un insumo valioso para los profesores de matemáticas, investigadores de su didáctica y encargados de elaborar libros de texto, interesados en diseñar propuestas de enseñanza en beneficio de mejorar la educación estocástica escolar.

Referencias

Alvarado, H. y Batanero, C. (2008). El significado del teorema central del límite en textos universitarios de probabilidad y estadística. *Estudios Pedagógicos, 34*(2),7-28.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L. y Spangler, D. (2020). *Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II* (GAISE II). American Statistical Association.

Batanero, C., Gea, M., Díaz- Levicoy, D. y Cañadas, G. (2015). Objetos matemáticos ligados a la regresión en los textos españoles de bachillerato. *Educación Matemática*, *27*(2), 9-35.

Bizet, V. Molina-Portillo, E. y Contreras, J.M. (2023a). Situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad según libros de texto chilenos. *Profesorado, revista de currículum y formación de profesorado*, *27*(2), 351-382.

Bizet, V., Molina-Portillo, E., Ruz, F. y Contreras, J.M. (2023b). Elaboración de una guía de Situaciones-problema sobre variable Aleatoria y sus aplicaciones a partir del currículo escolar chileno. *Educación Matemática. 35*(1), 163-190.

Bizet, V. y Ramos, E. (2022). Valoración de una situación didáctica para la enseñanza de variable aleatoria y distribución de probabilidad en la educación secundaria chilena. *Innovaciones Educativas*, *24*(36), 21–36.

Blanco, M., Bozt, J., Calderón, F., Jiménez, M., González, M., López, G., Romero, P., Díaz, M., Muñoz, G. y Rupin, P. (2009). *Matemática 2 proyecto bicentenario*. Santillana.

Blanco, M., Bozt, J., Calderón, F., Romero, L., Jiménez, L. y Jammet, C. (2009). *Matemática 3 proyecto bicentenario.* Santillana.

Choo-Kim, T. y Choo-Peng, T. (2015). Effects of the handheld technology instructional approach on performances of students of different achievement levels. *Computers and Education*, *82*, 306-314.

Cobo, B. (2003). *Significado de las medidas de posición central para los estudiantes de secundaria* [Tesis de Doctorado, Universidad de Granada].

Departamento de Investigaciones Educativas. (2014). *Matemática 4 proyecto bicentenario*. Santillana.

Díaz, E., Ortiz, N., Morales, K., Rebolledo, M., Barrera, R. y Norambuena, P. (2020). *Texto del estudiante de matemática 2° medio*. Ediciones SM.

Doukhan, C. y Gueudet, G. (2019). Students' difficulties at the secondary-tertiary transition: the case of random variables. En U. Jankvist, M. van den Heuvel-Panhuizen, y M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 2464-2471). Freudenthal Group y Freudenthal Institute of the Utrecht University y ERME.

Escolano, A. (2009). El manual escolar y la cultura profesional de los docentes. *Tendencias pedagógicas, 14,* 169-180.

Font, V. y Godino, J. (2006). La noción de configuración epistémica como herramienta de análisis de textos matemáticos: su uso en la formación de profesores. *Educación Matemática Pesquisa*, *8*(1), 67-98.

Font, V., Godino, J. y Gallardo, J. (2013). The emergence of objects from mathematical practices. *Educational Studies in Mathematics, 82*(1), 97-124.

Godino, J. (2017). Construyendo un sistema modular e inclusivo de herramientas teóricas para la educación matemática. En J. Contreras, P. Arteaga, G. Cañadas, M. Gea, B. Giacomone y M. López-Martín (Eds.), *Actas del Segundo Congreso International Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos* (pp. 1-20). Universidad de Granada.

Godino, J., Font, V. y Batanero, B. (2020). El enfoque ontosemiótico: implicaciones sobre el carácter prescriptivo de la didáctica. *Revista chilena de educación matemática, 12*(2), 3-15.

Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la investigación*. Editorial McGraw Hill Education.

Krippendorff, K. (1990). *Metodología de análisis de contenido: teoría y práctica*. Paidós.

Li, J., Cheng, J., An, T. y Zhou, D. (2021). Comparative Study on Binomial Distribution Content in High School Textbooks. En J. Wang (Ed.), *School mathematics textbooks in China, comparative studies and beyond* (pp. 107-136). Word scientific.

Lugo-Armenta, J. y Pino-Fan, L. (2021). Niveles de Razonamiento Inferencial para el Estadístico t-Student. *Bolema, 35*(71), 1776-1802.

Ministerio de Educación de Chile. (2015). *Bases Curriculares 7° básico a 2° medio. Gobierno de Chile.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2016). *Matemática, programa de estudio segundo medio.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019a). Chile. *Bases Curriculares 3º y 4º medio.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019b). *Programa de estudio matemática 3° medio para formación general*. Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019c). *Programa de estudio matemática 4° medio para formación general.* Unidad de Currículum y Evaluación.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standars for Mathematics.* NGACBP and CCSSO.

Ortiz, J. (2002). *La probabilidad en los libros de texto.* Grupo de Investigación en Educación Estadística de la Universidad de Granada.

Osorio, G., Norambuena, P., Romante, M., Gaete, D., Díaz, J., Celedón, J., Morales, K., Ortiz, N., Ramírez, P., Barrera, R. y Hurtado, Y. (2019). *Texto del estudiante de matemática 3° y 4° medio*. Ediciones SM.

Pfannkuch, M. (2018). International handbook of research in statistics education. En D. Ben-Zvi, K. Makar y J. Garfield (Eds.), *Reimagining curriculum approaches* (pp. 387-413). Springer.

Sánchez, E. y Landín, P. (2014). Levels of probabilistic reasoning of high school students about binomial problems. En E. Chernoff y B. Sriraman (Eds.), *Probabilistic thinking presenting plural perspectives* (pp. 581-597). Springer.

Setiawan, E. (2020). Introducing statistical inference to senior high school students: a textbook analysis. *Journal of Physics: Conference Series*, *1663*, 1-9.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*(1), 1-36.

Valdez, J. y Salinas, J. (2019). Análisis de las respuestas de estudiantes de bachillerato a problemas sobre la distribución normal. En J. M. Contreras, M. Gea, M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística* (pp. 1-10). Universidad de Granada.

Valverde, M. (2017). *Un estudio de la presentación de la distribución normal en los textos de bachillerato* [Tesis de Maestría, Universidad de Granada].

Vásquez, C. y Alsina, Á. (2015). Un modelo para el análisis de objetos matemáticos en libros de texto chilenos: situaciones problemáticas, lenguaje y conceptos sobre probabilidad. *Profesorado, revista de currículum y formación de profesorado, 19*(2), 441-462.

Vásquez, C. y Alsina. Á. (2017). Proposiciones, procedimientos y argumentos sobre probabilidad en libros de texto chilenos de educación primaria. *Profesorado, revista de currículum y formación de profesorado, 21*(1), 433-457.

CAPÍTULO 4: CONSTRUCCIÓN DE UN INSTRUMENTO Y APROXIMACIÓN A SU VALIDEZ

4.1 Introducción
4.2 **Estudio 4.** Elaboración de una guía de situaciones-problema sobre variable aleatoria y sus aplicaciones a partir del currículo escolar chileno
4.2.1 Introducción
4.2.2 Marco conceptual
4.2.3 Metodología
4.2.3.1 Contexto y materiales de la investigación
4.2.3.2 Fase 1: Construcción de guía de situaciones-problemas
4.2.3.3 Fase 2: Selección de un conjunto inicial de ítems a partir de la literatura
4.2.3.4 Fase 3: Selección de un conjunto final de ítem a partir de la validez de contenido por juicio de expertos
4.2.4 Resultados
4.2.4.1 Guía de situaciones-problemas
4.2.4.2 Análisis descriptivo de la validez de contenido por juicio de expertos
4.2.4.3 Estimación de la validez de contenido por juicio de expertos
4.2.5 Conclusiones
4.3 Análisis a priori de la versión inicial del instrumento
4.3.1 C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico
4.3.2 C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria
4.3.3 C-P_3 Reconocer la distribución de probabilidad de una variable aleatoria continua como una herramienta que permite ver la variación aleatoria
4.3.4 C-P_4 Establecer la función de distribución de una variable aleatoria discreta
4.3.5 C-P_5 Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria
distribución de sus probabilidades
4.3.6 C-P_6 La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real
4.3.7 C-P_7 La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real
4.3.8 C-P_8 Aproximar distribuciones de variables aleatorias discretas con un número grande de valores
4.4 **Estudio 5.** Cuestionario para valorar la comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar
4.4.1 Introducción
4.4.2 Antecedentes
4.4.3 Fundamentación
4.4.3.1 Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos
4.4.3.2 Campos de problemas de variable aleatoria y distribuciones de probabilidad
4.4.4 Metodología
4.4.4.1 Instrumento
4.4.4.2 Participantes
4.4.4.3 Procedimientos

4.1 Introducción

Este capítulo de la tesis doctoral presenta el proceso de construcción de un instrumento y aproximación a su validez de contenido y de constructo.

En el apartado 4.2 se exponen el Estudio 4 publicado en Educación Matemática, donde primero es abordado el diseño de una guía de situaciones-problemas sobre variable aleatoria y distribuciones binomial y normal, a partir del análisis de lineamientos curriculares chilenos y extranjeros (Franklin et al., 2005; NCTM, 2000; NGACBP y CCSSO, 2010). Luego se presenta la selección de un conjunto inicial de ítems representantes de las situaciones-problemas que componen la guía, desde la indagación de investigaciones previas y análisis de libros de texto. Después es realizada la validez de contenido de aquellos ítems por medio de la valoración de un grupo de expertos, doctores en educación matemática o psicología con experiencia en investigación en el área de didáctica de la estadística, donde su evaluación permite seleccionar un conjunto final de ítems. Además, se estima la validez de contenido de cada uno de ellos.

Posteriormente, en el apartado 4.3 es desarrollado el análisis a priori de la versión inicial del instrumento, allí son detallados los diversos objetos matemáticos primarios involucrados en las prácticas que permiten resolver correctamente cada ítem.

El apartado 4.4 presenta el Estudio 5 que está en evaluación en una revista sobre investigaciones en educación. En este estudio se lleva a cabo la validez de constructo, donde mediante un análisis factorial exploratorio es indagada la estructura de la versión inicial del instrumento y por medio de un análisis factorial confirmatorio se realiza una aproximación a confirmar la estructura de la versión final del instrumento. También es analizada la fiabilidad de la versión final del instrumento por medio del método de consistencia interna.

En el siguiente apartado se expone el índice de dificultad y discriminación de cada ítem que compone la versión final del cuestionario. Finalmente, el apartado 4.6 presenta el significado institucional evaluado en aquel instrumento.

4.2 Estudio 4[5]. Elaboración de una guía de situaciones-problema sobre variable aleatoria y sus aplicaciones a partir del currículo escolar chileno

Bizet, V., Molina-Portillo, E., Ruz, F. y Contreras, J.M. (2023c). Elaboración de una guía de situaciones-problema sobre variable aleatoria y sus aplicaciones a partir del currículo escolar chileno. *Educación Matemática, 35*(1), 163-190.

Resumen

Esta investigación tiene por objetivo proponer una guía que integre las situaciones-problemas relativas a la comprensión de variable aleatoria y sus aplicaciones en distribuciones de probabilidad promovidos por el currículo escolar chileno, fundamentada en las ideas de Transposición Didáctica y objetos matemáticos primarios del EOS (Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos). Primero, por medio del análisis de contenido a lineamientos curriculares chilenos e internacionales, fue creada la Guía de Situaciones-Problemas sobre Variable Aleatoria y sus Aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar Chileno (GSP-VADP). Posteriormente, a partir de la literatura ha sido seleccionado un conjunto inicial de ítems representantes de las situaciones-problemas que componen la guía, y luego a través de un juicio de expertos fue identificado el conjunto final. Los resultados muestran que la GSP-VADP está constituida por 33 situaciones-problemas, de aquellas emergen los restantes objetos matemáticos primarios. Esta herramienta es viable utilizarla para identificar ítems representantes de sus situaciones-problemas por medio de juicio de expertos, debido a que el conjunto final de ítems obtuvo un coeficiente de validez y concordancia bueno (0,87).

[5] Para evitar confusión y repetición en la numeración, se ha modificado la relativa a los epígrafes pertenecientes a los estudios que forman parte del compendio, haciendo referencia al capítulo en que se ubica cada estudio. También se ha cambiado la numeración de Tablas y Figuras con respecto a la versión publicada del artículo, haciendo referencia al estudio en que se enmarcan.

Palabras clave: variable aleatoria, distribuciones de probabilidad, situaciones-problemas, transposición didáctica, currículo escolar.

Abstract

This research has the aim of proposing a guide that integrate the problems-situations regarding the understanding of random variable and their applications in probability distributions promoted by the Chilean School Curriculum, based on the ideas of Didactic Transposition and OSA's primary mathematical objects (Ontosemiotic Approach to Mathematical Cognition and Instruction). First, through the content analysis to Chilean and international curricular documents, the Guide of Problems-Situations about Random Variable and their Applications in Probability Distributions according to the School Chilean Curriculum (GPS-RVPD) was created. Subsequently, from the literature has been selected an initial item group that is a representation of the problems-situations that compose the guide, and then through an expert judgment the final group was identified. The results show that the GPS-RVPD is constituted by 33 problems-situations, from which emerge the rest primary mathematical objects. This tool is feasible to use for identifying items representing their problems-situations, through expert judgment, this is because the final item group obtained a good validity and concordance coefficient (0,87).

Keywords: random variable, probability distribution, problems-situations, didactic transposition, school curriculum.

4.2.1 Introducción

La variable aleatoria y las distribuciones de probabilidad son tópicos esenciales en la educación estocástica dada su importancia en la teoría de la probabilidad y en la práctica estadística (Heitele, 1975; Batanero, 2004). Por esta razón se les da un lugar importante a estos temas en el currículo de Chile (Ministerio de Educación de Chile [MINEDUC], 2015; MINEDUC, 2019a) y en el currículo de la mayoría de los países, por ejemplo, el de Estados Unidos (National Council of Teachers of Mathematics [NCTM], 2000; National Governors Association Center for Best Practices [NGACBP] y Council of Chief State School Officers [CCSSO], 2010).

En las investigaciones que tienen como objeto de estudio los libros de texto de probabilidad y estadística, se ha constatado la escasez de situaciones-problemas sobre el tema de variable aleatoria y distribuciones (Ortiz, 2002, Valverde, 2017). Por otro lado, a pesar de que la variable aleatoria es inseparable de su distribución de probabilidad (Ruiz, 2013), se ha notado que la literatura previa ha sido persistente en reportar deficiencias en los conocimientos utilizados en el proceso de enseñanza-aprendizaje de la distribución binomial a nivel escolar (Sánchez y Landín, 2014), y la variable aleatoria (Ruiz, 2006) o la distribución normal en el ámbito universitario (Tauber, 2001).

Por ello, identificamos una problemática vigente en torno a la carencia de herramientas que permitan valorar de manera articulada y simultánea la comprensión de la variable aleatoria y sus aplicaciones en las distribuciones de probabilidad binomial y normal. Lo anterior se agudiza frente al inminente tratamiento de este contenido durante la transición de egreso de la educación escolar e inicio de la educación superior, que se espera sea coherente con los lineamientos curriculares actuales. Por tanto, en este trabajo se ha propuesto abordar la problemática descrita previamente, para lo cual se establece el objetivo de proponer una guía que integre las situaciones-problemas relativas a la comprensión de variable aleatoria y sus aplicaciones en distribuciones de probabilidad, promovidas por el currículo escolar chileno, donde las situaciones-problemas hacen referencia a tareas, actividades o ejercicios que originan la actividad matemática (Godino, 2002).

4.2.2 Marco conceptual

La investigación se fundamenta en la noción de tansposición didáctica, entendida como el cambio realizado al conocimiento científico (saber sabio) para poder ser enseñalado (Chevallard, 1980). Dentro de esta se distinguen dos tipos (Chevallard, 1991): la transposición didáctica externa, que corresponde a la transformación realizada del conocimiento científico (saber sabio) al contenido expuesto en el curriculo escolar (saber a enseñar); transposición interna, que hace referencia al cambio del contenido presente en los lineamientos curriculares (saber a enseñar) en una forma de conocimiento que sea accesible a los estudiantes (saber enseñado).

Además, en la transposición didáctica intervienen agentes, tales como (Chevalard, 1991): (i) expertos, quienes vigilan en el proceso de transformación que el saber enseñado no sea una distorción del saber sabio; (ii) documentos curriculares, los cuales exponen los contenidos de referencia sobre los que debe actuar el profesor; (iii) profesor, quien es uno de los responsables del proceso de transposición didáctica, y (iv) estudiantes, quienes son los receptores de aquel producto.

También este estudio se sustenta en elementos del Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos (EOS) (Godino et al., 2007), teoría que propone que el origen de la actividad matemática son las situaciones-problemas, es decir, tareas, actividades o ejercicios (Godino, 2002), y para su resolución se realizan determinadas prácticas matemáticas. Aquellas prácticas son específicas actuaciones y expresiones verbales o gráficas que pueden ser compartidas en una institución (llamado significado institucional) o realizadas por una persona (denominado significado personal) para resolver una tarea o comunicar su solución (Godino, 1994).

En el EOS los objetos corresponden a todo aquello a lo que se hace referencia en la práctica matemática, por lo que según la función que cumplan es posible clasificarlos en la siguiente tipología de objetos matemáticos primarios (Godino, 2002): a) situaciones-problemas (S-P), tareas, actividades o ejercicios que promueven la actividad matemática; b) lenguaje (L), términos matemáticos en sus diversos registros de representación (verbal, simbólico, tabular y gráfico) utilizados para enunciar o resolver tareas; c) conceptos (C), descripciones o definiciones relacionadas a un objeto matemático, aplicadas en la resolución de actividades; d) proposiciones (PP), características o propiedades de conceptos, usadas en solucionar tareas; e) procedimientos (P), algoritmos, técnicas de cálculos u operaciones desarrollados para responder actividades; y f) argumentos (A), enunciados empleados para aprobar o justificar proposiciones y procedimientos, o la respuesta a tareas. Así, un conjunto de situaciones-problemas vinculadas recíprocamente, que comparten sus representaciones, procesos o soluciones similares constituyen un campo de problema (C-P) (Godino, 1999).

De esta manera, el EOS se ha utilizado en la determinación de los campos de problemas que estructuran la guía diseñada, a partir de la indagación de investigaciones previas en torno a la epistemología de los temas en cuestión. La noción de transposición didáctica interna fue empleada en el análisis de documentos curriculares tanto chilenos como internacionales y proceso de identificación, clasificación y reducción de unidades de análisis. También el Enfoque Ontosemiótico se ha usado en el proceso de inferencia de situaciones-problemas a partir de las unidades de análisis, además en la selección desde la literatura de tareas, actividades o ejercicios que representan aquellas situaciones-problemas que componen la guía.

4.2.3 Metodología

El estudio ha sido desarrollado desde un enfoque mixto y es de tipo exploratorio-descriptivo (Hernández et al., 2014). Fue utilizado el diseño transeccional o transversal (Hernández et al., 2014) debido a que se recopilaron datos (empíricos) en un solo momento, obtenidos de las opiniones de expertos iberoamericanos que validaron la guía diseñada.

4.2.3.1 Contexto y materiales de la investigación

El contexto de esta investigación ha sido la educación escolar chilena, por consiguiente, en primer lugar, se indagó el principal documento curricular del país propuesto por el Ministerio de Educación, denominado Bases Curriculares (MINEDUC, 2015; MINEDUC, 2019a) únicamente las destinadas a la educación secundaria (grado 9 hasta 12). Además, se incluyó el análisis de los Programas de Estudio de Matemática (MINEDUC, 2016; MINEDUC, 2019b; MINEDUC 2019c), pues es un instrumento que hace posible la ejecución de las Bases Curriculares. En estos documentos fueron estudiados los Objetivos de Aprendizaje (O.A.) e Indicadores de Evaluación (I. Ev.) sobre tópicos estocásticos.

La elaboración de las Bases Curriculares, se sustentan en diversas fuentes, una de ellas es la revisión de lineamientos curriculares internacionales principalmente de países pertenecientes a la Organización para la Cooperación y el Desarrollo Económico (MINEDUC, 2009). En consecuencia, en un segundo momento se analizaron los siguientes documentos:

- Los estándares americanos como los Principios y Estándares para la Educación Matemática (NCTM, 2000) y los Estándares Estatales Básicos Comunes para las Matemáticas (NGACBP y CCSSO, 2010), dirigidos a la educación preescolar y escolar (grados K al 12). En el primero de ellos se han estudiado las Expectativas de Aprendizaje (E.A.) del estándar de Análisis de Datos y Probabilidad, mientras que en el segundo los Dominios (D) y Estándares (E) pertenecientes a la categoría conceptual de Estadística y Probabilidad.

- Los Lineamientos para la Evaluación y Enseñanza en Educación Estadística, conocido como proyecto GAISE (Franklin et al., 2005) que abarca desde la educación preescolar a grado 12, pues es un documento de referencia internacional para afrontar los desafíos de la educación estadística a nivel escolar. En él fue estudiado el apartado sobre el rol de la probabilidad en estadística.

La investigación se organizó en tres fases, cada una es detalla a continuación.

4.2.3.2 Fase 1: Construcción de guía de situaciones-problemas

En la primera fase de la investigación, se empleó la técnica de análisis de contenido (Krippendorff, 1990). Esta es una técnica para analizar de forma sistemática documentos, con el propósito de realizar inferencia reproducible e identificar en sus fragmentos la presencia o ausencia de alguna característica del tema. Por ello se ha utilizado en la identificación y clasificación de normativas vigentes en torno a la enseñanza-aprendizaje de la variable aleatoria y sus aplicaciones en algunas distribuciones de probabilidad, organizadas en categorías, según los ocho campos de problemas abajo definidos.

Aquellos campos de problemas fueron establecidos en un proceso cíclico e inductivo por medio del análisis a la epistemología de la variable aleatoria (Amrani y Zaki, 2015; Dinges, 2005; Ruiz, 2013). También este aspecto ha sido indagado en la distribución binomial (García- García et al., 2022; Vilca, 2015) y distribución normal (Stahl 2006; Tauber, 2001). Así se obtuvieron los siguientes:

1. Identificar la variable aleatoria como una función presente en el contexto probabilístico.
2. Reconocer la distribución de probabilidad de una variable aleatoria discreta (función de cuantía) como una herramienta que permite ver la variación aleatoria.
3. Reconocer la distribución de probabilidad de una variable aleatoria continua (función de densidad) como una herramienta que permite ver la variación aleatoria.
4. Establecer la función de distribución de una variable aleatoria discreta.
5. Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria
6. La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real.
7. La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real.
8. Aproximar distribuciones de variables aleatorias discretas con un número grande de valores.

Las Unidades de Análisis (U.A.) fueron definidas a partir de la revisión de los documentos curriculares señalados y clasificadas en los campos de problemas anteriores, para ello se emplearon letras minúsculas del alfabeto latino (a, b, c, …, z). Luego por medio del método de Rivas (2014), las U.A. han sido comparadas y reducidas utilizado el comentario “contenida en”, debido a que algunas unidades estaban contenidas en otra o no entregaban nueva información, todo ello con el propósito de lograr representar su información en una U.A. final. Un ejemplo de este procedimiento se expone en la Tabla 4.1.

Tabla 4.1

Ejemplo del proceso de identificación, clasificación y reducción de las Unidades de Análisis

Campo de problema	Unidades de Análisis
2. Reconocer la distribución de probabilidad de una variable aleatoria discreta (o función de cuantía) como una herramienta que permite ver la variación aleatoria	MINEDUC (2015, p. 124): OA 10 Mostrar que comprenden las variables aleatorias finitas a. Calculando su probabilidad. *U.A. a contenida en h* b. Graficando sus distribuciones. *U.A. b contenida en d* MINEDUC (2016, p. 154): I. Ev. c. Determinan las probabilidades de una variable aleatoria aplicando la terminología P(X = xi). *U.A. c contenida en h* d. Elaboran tablas y gráficos para representar la distribución de una variable aleatoria finita. *U.A. d* NCTM (2000, p. 324): E. A. e. Comprender los conceptos de espacio muestral y distribución de probabilidad y construir espacios muestrales y distribuciones en casos simples. *U.A. e contenida en h* f. Usar simulaciones para construir distribuciones de probabilidad empíricas. *U.A. f contenida en i* NGACBP y CCSSO (2010, p.82-83): D. Usar la probabilidad para tomar decisiones g. E.1: Graficar la distribución de probabilidad correspondiente utilizando el mismo gráfico como para distribuciones de datos. *U.A. g contenida en d* h. E.3: Desarrollar una distribución de probabilidad para una variable aleatoria definida en un espacio muestral en el que se puede calcular las probabilidades teóricas. *U.A. h* i. E.4: Desarrollar una distribución de probabilidad para una variable aleatoria definida en un espacio muestral en el que las probabilidades se asignan empíricamente. *U.A. i* Franklin et al. (2005): El rol de la probabilidad en estadística j. Se requiere la comprensión de conceptos probabilísticos, como las distribuciones de probabilidad. *U.A. j contenida en h*

Nota. Elaboración propia.

Posteriormente, a partir de las Unidades de Análisis final se procedió a inferir situaciones-problemas (S-P) alusivas a la comprensión de la variable aleatoria y distribuciones de probabilidad, donde se presentaron los siguientes casos: (i) una U.A. final originó una única situación-problema; (ii) una U.A final originó dos o tres situaciones-problemas, ejemplo de ello se muestra en la Tabla 4.2.

Tabla 4.2

Ejemplo del proceso de inferencia de situaciones-problemas

Unidades de Análisis final	Situaciones-Problemas
U.A. final i. Desarrollar una distribución de probabilidad para una variable aleatoria definida en un espacio muestral en el que las probabilidades se asignan empíricamente (NGACBP; CCSSO, 2010, p.82-83)	2.1 Determinar las probabilidades asociada a los valores de una variable aleatoria discreta desde un enfoque frecuencial
U.A. final h. Desarrollar una distribución de probabilidad para una variable aleatoria definida en un espacio muestral en el que se puede calcular las probabilidades teóricas (NGACBP; CCSSO, 2010, p.82-83)	2.2 Determinar las probabilidades asociada a los valores de una variable aleatoria discreta desde un enfoque clásico 2.3 Definir la función de probabilidad de una variable aleatoria utilizando la terminología $P(X = x_i)$.
U.A. final d. Elaboran tablas y gráficos para representar la distribución de una variable aleatoria finita (MINEDUC, 2016, p. 154	2.4 Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta 2.5 Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta

Nota. Elaboración propia.

Finalmente, como resultado de esta fase se obtuvo un conjunto de 33 situaciones-problemas, que sintetizaron los objetivos de aprendizaje e indicadores de evaluación en torno a la comprensión de variable aleatoria y sus aplicaciones en el modelo binomial y normal, presentes en documentos curriculares escolares. En su conjunto conformaron la Guía de Situaciones-Problemas sobre Variable Aleatoria y sus Aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar Chileno (GSP-VADP).

4.2.3.3 Fase 2: Selección de un conjunto inicial de ítems a partir de la literatura

En esta fase se seleccionó de la literatura actividades, tareas o ejercicios concretos, representantes de las situaciones-problemas que componen la guía sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad. La búsqueda se realizó en estudios, tanto donde se hayan utilizado cuestionarios para evaluar la comprensión de la variable aleatoria y/o distribución binomial o normal, como en aquellos que se presentan propuestas de enseñanza en torno a los conceptos en cuestión. En el caso que, en las investigaciones indagadas no se encontró una pregunta pertinente para valorar una situación-problema, se procedió inicialmente a buscar una interrogante en libros de Textos Escolares de Matemática (chilenos) y en última instancia esta fue creada según el contexto de algún enunciado propuesto para evaluar otra situación-problema.

Para cada situación-problema que integra la guía, fue preparado un trío de ítems diferentes, aunque en algunos casos un ítem valoró más de una situación-problema. De esta manera se obtuvo un primer conjunto constituido por 91 ítems, la mayor parte fueron preguntas de tipo abierta, debido a que entregan información más detallada (Hernández et al., 2014).

4.2.3.4 Fase 3: Selección de un conjunto final de ítem a partir de la validez de contenido por juicio de expertos

En este momento se procedió a la validez de contenido de los ítems por medio de la valoración por juicio de expertos, su evaluación permitió seleccionar de cada trío propuesto al ítem mejor calificado para el conjunto final. Participaron seis expertos de diferentes nacionalidades (Argentina, Chile, España, México y Portugal), dado que se recomienda considerar entre cinco a diez (Barraza, 2007), todos ellos doctores en educación matemática o psicología con una reconocida experiencia en investigación en el área de didáctica de la estadística. A cada uno de los expertos se les envió, por medio de correo electrónico, una invitación en la cual se especificó el contexto de la investigación y objetivo del instrumento y se solicitó su colaboración para valorar: el grado en que cada ítem es adecuado para evaluar la situación-problema donde fue clasificado; elegir aquel ítem que considera que evalúa de mejor manera la situación-problema, entre los tres que se proponen por cada una. Para ello fue propuesta una escala Likert, de 1 (nada relevante) a 5 (muy relevante), sumado a un apartado de sugerencias donde pudieran recomendar mejoras en la redacción de los ítems.

Posteriormente se realizó un análisis descriptivo de los resultados por medio de indicadores estadísticos sobre el centro (media: $\overline{\mathrm{X}}$ y mediana: M_e) y la dispersión (desviación típica) de las valoraciones de los expertos. Así de los tres ítems propuestos para cada situación-problema, ha sido seleccionado aquel con media más alta. Para los casos donde existieron dos ítems con igual media, fue elegido aquel que tiene una mediana con mayor valor, mientras que, si hubo dos ítems con igual media y mediana, fue seleccionado el con menor desviación típica.

Finalmente, para analizar el grado en que cada ítem seleccionado se ajustó a una situación-problema, fue calculado su Coeficiente de Validez de Contenido (CVC: Hernández-Nieto, 2002), el cual permite valorar el grado de acuerdo de los expertos respecto a cada uno de los diferentes ítems y al instrumento en general. Para interpretar el CVC de un ítem se tuvo en cuenta la siguiente escala: (i) ítem inaceptable $(CVC < 0{,}60)$; ítem deficiente $(0{,}60 \leq CVC \leq 0{,}70)$; (iii) ítem adecuado $(0{,}70 < CVC \leq 0{,}80)$; (iv) ítem bueno $(0{,}80 < CVC \leq 0{,}90)$; (v) ítem excelente $(0{,}90 < CVC)$. De esta manera se consideraron aceptables los ítems con un coeficiente de validez y concordancia mayor que 0,70.

4.2.4 Resultados

4.2.4.1 Guía de situaciones-problemas

La Tabla 4.3 presenta la Guía de Situaciones-Problemas sobre Variable Aleatoria y sus Aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar Chileno (GSP-VADP). Como se puede apreciar el primer campo de problema, constituido por cuatro situaciones-problemas, aborda el reconocimiento de la variable aleatoria ligada a un fenómeno aleatorio, dado que es el primer paso para estudiar distribuciones de probabilidad. El segundo y tercer campo de problema, compuesto por cinco y tres situaciones-problemas respectivamente, tiene relación con el establecimiento de la función de probabilidad: en una primera instancia para la variable aleatoria de tipo discreta, que permite determinar la probabilidad de que esta tome valores aislados; en un segundo momento para la variable aleatoria de tipo continua que, hace posible calcular la probabilidad de que esta tome valores en un intervalo de la recta real. El cuarto campo de problema integrado por tres situaciones-problemas, hace referencia a estudiar la función de distribución de una variable aleatoria discreta, dado que son tópicos inseparables y esta función permite representar la situación aleatoria en que ambas están involucradas. El quinto campo de problema, conformado por seis situaciones-problemas, aborda algunos valores de posición central o de dispersión vinculados a la variable aleatoria.

El sexto y séptimo campo de problema, constituidos cada uno por cinco situaciones-problemas, trabajan la aplicación de la variable aleatoria en los modelos de probabilidad, para las de tipo discreta (modelo binomial) y continuas (modelo normal) correspondientemente. El octavo y último campo de problema, constituido por dos situaciones-problemas, está relacionado con la aproximación de la distribución binomial por la normal.

Por tanto, a partir de las 33 situaciones-problemas identificadas emergen los restantes tipos de objetos matemáticos (lenguaje, conceptos, proposiciones, procedimientos y argumentos), y que en su conjunto reflejan el significado institucional (de referencia) de la variable aleatoria y sus aplicaciones en distribuciones de probabilidad en el contexto escolar chileno.

Tabla 4.3

GSP-VADP

Campos de problemas	Situaciones-Problemas
1. Identificar la variable aleatoria como una función presente en el contexto probabilístico	1.1 Diferenciar entre variables aleatorias y variables con dependencia funcional 1.2 Definir variables aleatorias finitas en experimentos aleatorios 1.3 Identificar dominio de una variable aleatoria finita 1.4 Identificar recorrido o valores de una variable aleatoria finita
2. Reconocer la distribución de probabilidad de una variable aleatoria discreta (o función de cuantía) como una herramienta que permite ver la variación aleatoria	2.1 Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial 2.2 Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico 2.3 Definir la función de probabilidad de una variable aleatoria utilizando la terminología $P(X = x_i)$ 2.4 Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta 2.5 Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta
3. Reconocer la distribución de probabilidad de una variable aleatoria continua (o función de densidad) como una herramienta que permite ver la variación aleatoria	3.1 Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua 3.2 Calcular algunas probabilidades asociadas a una variable aleatoria continua 3.3 Determinar si la función dada es una función de probabilidad de una variable aleatoria continua
4. Establecer la función de distribución de una variable aleatoria discreta	4.1 Determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta 4.2 Definir la función de distribución (o función de probabilidad acumulada) de una variable aleatoria discreta 4.3 Representar en lenguaje gráfico la función de distribución (o función de probabilidad acumulada) de una variable aleatoria discreta
5. Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria	5.1 Calcular la media o valor esperado de una variable aleatoria discreta 5.2 Interpretar la media o valor esperado de una variable aleatoria discreta 5.3 Calcular la varianza de una variable aleatoria discreta 5.4 Calcular la desviación estándar de una variable aleatoria discreta 5.5 Interpretar la desviación estándar de una variable aleatoria discreta 5.6 Identificar la media y desviación estándar de una variable aleatoria continua
6. La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real	6.1 Identificar situaciones que pueden modelizarse a través de una distribución binomial 6.2 Determinar los parámetros asociados a una distribución binomial como n (número de ensayos), p (probabilidad de éxito) y q (probabilidad de fracaso) 6.3 Calcular la media, varianza y desviación estándar de una distribución binomial 6.4 Determinar probabilidades asociadas a una distribución binomial de forma manual 6.5 Determinar probabilidades asociadas a una distribución binomial empleando una herramienta tecnológica
7. La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real	7.1 Identificar situaciones que pueden modelizarse a través de una distribución normal 7.2 Determinar probabilidades asociadas a una distribución normal de forma manual

	7.3 Describir la tendencia de los datos representados gráficamente empleando la aproximación de la binomial por la normal 7.4 Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal 7.5 Determinar probabilidades asociadas a una distribución normal empleando una herramienta tecnológica
8. Aproximar distribuciones de variables aleatorias discretas con un número grande de valores	8.1 Calcular los parámetros asociados a una distribución normal como μ (media) y σ (desviación estándar) 8.2 Calcular probabilidades en el contexto de aproximación de la distribución binomial a la normal

Nota. Elaboración propia.

A continuación, se muestra la selección de cada ítem representante de una situación-problema a través de la valoración por juicio de expertos. Es importante mencionar que, en las fases de selección de ítems a partir de la literatura y elección según juicio de expertos, se han excluido dos situaciones-problemas (6.5 y 7.5), debido a que se proyecta que el conjunto final de ítems sea aplicado a estudiantes individualmente y en forma presencial o en línea. Además, en las fases señalas se incluyeron ítems para valorar dos situaciones-problemas (6.1 y 7.1) fundamentadas en la alfabetización estadística en términos de Gal (2002), con el propósito de aportar con información sólida y empírica a la investigación de esta área.

4.2.4.2 Análisis descriptivo de la validez de contenido por juicio de expertos

A los seis expertos participantes se les presentó un documento donde la primera sección contenía tres ítems para cada una de las 31 situaciones-problemas previamente identificadas. Su valoración hizo posible seleccionar de cada trío de ítems propuesto, aquel que fue más relevante para ellos (ítem sombreado), a través del mayor valor medio dado a cada uno. Por ejemplo, para la situación-problema 1.1 fueron propuestos los ítems 1, 2A y 3.1, por lo que ha sido seleccionado el ítem 1, cuya media es de 4,33 puntos. La Tabla 4.4, muestra los siguientes casos presentados en la elección de ítems que componen la versión definitiva de esta sección:

i. En tres situaciones-problemas (1.4, 6.2 y 7.3), un par de ítems fueron evaluados con la misma puntuación media, seleccionándose aquel con mayor mediana. En concreto, para la situación problema 1.4, los ítems 1 y 3.2C, poseen igual media (4 puntos), entonces fue seleccionado el ítem 3.2C, cuya mediana es 4,5 puntos.

ii. En dos situaciones-problemas (5.2 y 7.1), un par de ítems han sido valorados con igual media y mediana, entonces fue elegido el que presenta menor desviación típica. En particular, para la situación-problema 5.2, los ítems 14.1A y 15C tienen misma media (4 puntos) y mediana (4 puntos), así fue elegido el ítem 15C cuya desviación típica es 0 punto.

iii. En la situación-problema 4.2, los ítems 8A y 9A fueron calificados con igual media (4,683 puntos), mediana (4,7 puntos) y desviación típica (0,37 puntos). Por tanto, ha sido escogido el ítem 8A, ya que posee igual contexto que los ítems seleccionados para la situación-problema 4.1 y 4.3.

Tabla 4.4

Selección de ítems y resumen estadístico de puntuaciones de expertos a los ítems

Situación-Problema	Ítem	$\overline{X}$	M_e	Desv. típ.	Situación-Problema	Ítem	$\overline{X}$	M_e	Desv. típ.
1. 1	*1*	*4,33*	5	1,21	5.2	13.1B	3,8	4,4	1,47
	2A	3,83	4	1,17		14.1A	4	4	0,89
	3.1	2,33	2	1,51		*15C*	*4*	*4*	*0*
1. 2	*1*	*4,33*	4,5	0,82	5.3	13.2A	4	4	0,89
	2C	3,83	4	1,17		14.2A	3,67	4	1,03
	3.2 B	4	4,5	1,55		*15B*	*4,5*	4,5	0,55
1.3	1	3,17	3	1,17	5.4	13.2A	4	4	0,89
	2B	*4,17*	4,5	0,98		*14.2A*	*4,33*	4,5	0,82
	3.2 A y C	4	4	1,1		15B	4,17	4	0,75
1.4	1	4	4	0,89	5.5	13.2B	3,83	4	1,17
	2B	3,67	3,5	1,21		14.2B	3,67	3,5	1,21
	3.2 C	*4*	*4,5*	1,27		*15.C*	*4*	4	1,1
2.1	4.2	3	2,5	1,67	5.6	16	2,67	3	1,51
	5.1 A	*3,83*	4	1,17		*17*	*3,33*	4	1,51
	6.2	3,67	4	1,03		18	2	1	1,55
2.2	4.1 A	3,8	3,9	1,17	6.1	19.1	3,83	4	1,33
	5.1 B	3	3	1,1		*20*	*4,17*	4	0,75
	6.1 A	*4*	4	1,1		21	3,83	4	1,17
2.3	4.1 B	3	3	0,89	6.2	19.1	3,5	3,5	1,05
	5.2	3,83	4,5	1,47		*20*	*4,17*	*4,5*	0,98
	6.1 B	*4,17*	4	0,75		21A y B	4,17	4	0,75
2.4	4.1 B	3,33	3,5	1,21	6.3	19.2A y B	4,17	4,5	1,17
	5.1 B	*4,17*	4	0,75		*20 E y F*	*4,6*	4,8	0,49
	6.1 B	3,83	4	0,98		21C y D	4,5	5	0,84
2.5	4.1 C	4,17	4,5	0,98	6.4	19.1	4,17	4,5	1,17
	5.1 C	4,17	4,5	0,98		*20A, B, C y D*	*4,5*	4,5	0,55
	6.1 C	*4,33*	4,5	0,82		21B	4,33	4,5	0,82
3.1	*10A*	*3,83*	4	0,98	7.1	22	4,17	4,5	1,17
	11A	3,5	3,5	1,38		*23*	*4,17*	*4,5*	*0,98*
	12.1 B	3,33	3,5	1,63		24	3	3,5	1,27
3.2	10C y D	3,67	4	0,52	7.2	25	4,17	4	0,75
	11B	*3,83*	*4*	0,98		*26*	*4,33*	4,5	0,82
	12.2 y 12.3	3,67	4	1,37		27	4	4	1,1
3.3	10B	4	4	0,63	7.3	28	2,83	2,5	1,47
	11B	*4,17*	4	0,75		*29*	*4,5*	*5*	0,84
	12.1 A	4	4	1,1		30	4,5	4,5	0,55
4.1	7A	3,67	4,5	1,75	7.4	31	3,17	3,5	0,98
	8A	*4*	4	0,89		*32*	*3,83*	4	0,75
	9C	3,83	4	1,17		33	3,17	3	0,75
4.2	7B	4	4	0,89	8.1	34A	3,83	4	0,75
	8A	*4,683*	*4,7*	*0,37*		*35A*	*4*	4	0,63

	9A	4,683	4,7	0,37		36	3,67	3,5	0,82
4.3	7B	3,83	4	1,17	8.2	*34B*	*4,17*	4	0,75
	8B	*4,5*	4,5	0,55		35B	4	4	0,89
	9B	4,33	4,5	0,82		36	3,83	3,5	0,98
5.1	*13.1B*	*4,33*	4,5	0,82					
	14.1 A y B	4,17	4	0,75					
	15A	3,67	4	1,03					

Nota. Elaboración propia.

Como resultado se puede apreciar que de los 31 ítems elegidos para la versión final de esta sección del documento: 20 de ellos fueron muy relevantes para los expertos, dado que cada uno fue calificado en promedio con más de 4 puntos; seis han sido relevantes pues, cada uno fue evaluado en promedio con 4 puntos; cinco fueron medianamente relevantes ya que, cada uno ha sido valorado en promedio con más de 3 puntos, pero menos que 4.

Además, según se expone en la Tabla 4.4 los expertos han valorado positivamente las situaciones-problemas que componen cada campo de problema. Una evidencia de ello es que, en cada uno de los 31 ítems escogidos para la versión final de la primera sección del documento, su evaluación obtuvo una mediana igual o mayor a 4 puntos. Concretamente: en cuatro de estos ítems la mediana de sus puntuaciones es mayor a 4,5 puntos, es decir, la mitad de los expertos valoraron cada interrogante con 5 puntos a diferencia de los demás evaluadores que les asignaron 5 o menos puntuación; en 12 de ellos, las puntuaciones evidenciaron una mediana de 4,5 puntos, lo cual significa que tres jueces evaluaron a cada ítem con 5 puntos y los tres participantes restantes lo han hecho con 5 o menor valor; en 15 de esos, la mediana de sus puntuaciones es 4, es decir, tres de los jueces valoraron a estos ítems con 4 o 5 puntos y la otra mitad de los participantes les han asignado un valor igual o menor a 4.

Por tanto, para los expertos participantes, cada ítem seleccionado es representativo de la situación-problema en la cual han sido propuesto, además de ser relevante y coherente para evaluar la situación-problema consultada. También estos ítems representan las tareas, ejercicios y actividades que promueve el currículo escolar chileno para valorar la comprensión de la variable aleatoria y distribuciones binomial y normal.

Respecto a la segunda sección del documento presentado a los expertos, donde se abordan situaciones-problemas acerca de distribuciones de probabilidad fundamentadas en la alfabetización estadística, la Tabla 4.5 contiene los resultados sobre su valoración. En ella ha sido posible observar que se presentó el caso que una de las situaciones-problemas (6.1) obtuvo dos ítems calificados con igual media, mediana y desviación típica. Por tanto, ha sido escogido aquel, en que el estudiante debe poner en juego habilidades como evaluar e interpretar la información estadística presente en su enunciado, más que realizar procedimientos (algoritmos, operaciones, etc.).

Tabla 4.5

Selección de ítems sobre alfabetización estadística y resumen estadístico de puntuaciones de expertos

Situación-Problema	Ítem	$\bar{X}$	M_e	Desv. típ.
6.1 Identificar situaciones que pueden modelizarse a través de la distribución binomial	*37*	*4,5*	*5*	*0,84*
	38	4	4	1,1
	39	4,5	5	0,84
7.1 Identificar situaciones que pueden modelizarse a través de la distribución normal	*40*	*4,7*	5	0,52
	41	4	4	1,1
	42	4	4	1,1

Nota. Elaboración propia.

También la Tabla 4.5 muestra que los dos ítems que fueron seleccionados para la versión definitiva de esta sección han sido evaluados por los participantes como muy relevantes, pues cada uno fue calificado en promedio con más de 4 puntos. Aún más, en aquellos ítems la mediana de sus puntuaciones es 5, lo cual significa que tres de los expertos asignaron a cada uno la máxima valoración (5) a diferencia de los otros tres que los calificaron con 5 o menos puntos.

De esta manera se obtuvo un conjunto final de ítems que representan concretamente gran parte de las situaciones-problemas de la GSP-VADP. En su resolución se manifiestan conceptos, lenguaje, procedimientos, propiedades y argumentos, que en conjunto caracterizan el significado institucional (evaluado) de la variable aleatoria y sus aplicaciones en las distribuciones de probabilidad a nivel escolar. La Tabla 4.6 presenta información relativa a las fuentes bibliográficas de aquellos ítems, organizada según la situación-problema que valora cada uno.

Tabla 4.6

Fuentes bibliográficas del conjunto final de ítems

Situación-Problema	Ítem y Fuente
1.1	1.1 Adaptación de Fernández et al. (2013)
1.2	
1.3	1.2 Flores et al. (2014)
1.4	1.3 Adaptación de Flores et al. (2014) y Salazar (2014)
2.1	2.1a Flores et al. (2014)
2.2	2.2a Salazar (2014)
5.4	5.2a Elaboración propia
5.5	5.2c Elaboración propia
5.6	10.a Elaboración propia y posee información de Ramírez (2018)
6.1	6. Alvarado y Retamal (2014) 9. Elaboración propia y posee información de Colomé (2015) *Alfabetización estadística

2.3	2.2b Salazar (2014)
2.4	2.1b Flores et al. (2014)
2.5	2.2c Elaboración propia
3.1	3.1 Muñoz et al. (2013)
3.2	3.2b Muñoz et al. (2013)
3.3	3.2a Adaptación de Muñoz et al. (2013)
4.1	4.a Elaboración propia
4.2	4.a Elaboración propia
4.3	4.b Chacón et al. (2018)
5.1	5.1 Guerrero et al. (2016)
5.2	5.2c Elaboración propia
5.3	5.2b Osorio et al. (2019)
6.2	6. Alvarado y Retamal (2014) y adaptación de Sánchez y Carrasco (2018)
6.3	6.e y 6.f Elaboración propia
6.4	6.a, 6b, 6c y 6d Alvarado y Retamal (2014) y adaptación de Sánchez y Carrasco (2018)
7.1	7.1 Morales et al. (2008) 10.b Elaboración propia y posee información de Ramírez (2018) *Alfabetización estadística
7.2	7.2 González y Ojeda (2017)
7.3	7.3 Adaptación de MINEDUC (2019c)
7.4	7.4 Tauber (2001)
8.1	8.1 Norambuena et al. (2019)
8.2	8.2 Alvarado y Batanero (2007)

Nota. Elaboración propia.

En la Tabla 4.7 es expuesto un ejemplo de ítems seleccionados para el conjunto final, los cuales representan las situaciones-problemas correspondientes al campo de problema dos. En este caso, es posible apreciar los elementos que consideraron significativos los expertos para elegir las situaciones-problemas relativas a la distribución de probabilidades de una variable aleatoria discreta. Respecto a la situación-problema 2.1 los expertos seleccionaron el ítem 2.1A. Por tanto, se puede reconocer que consideraron relevante el proponer una situación real y cercana a estudiantes, la cual admite una estrategia de resolución por tanteo, donde para entender y dar respuesta a la tarea planteada se pueden llevar a cabo en repetidas ocasiones el experimento aleatorio involucrado.

Tabla 4.7

Ejemplo de ítems seleccionados para su conjunto final

<table>
<tr><th>Campo de problema 2: Reconocer la distribución de probabilidad de una variable aleatoria discreta (o función de cuantía) como una herramienta que permite ver la variación aleatoria</th><th>Situación-problema</th></tr>
<tr><td>2.1 (Flores, García y Sánchez, 2014) Llamaremos lanzamiento a la acción de lanzar tres monedas al aire al mismo tiempo (de preferencia de la misma denominación). Ahora imagínate que se realizan 1000 lanzamientos y en cada uno de ellos se observa la variable “número de caras que ocurren”.
A) En la siguiente tabla anota en la fila de arriba, los posibles valores de la variable y en la fila de abajo, el número de veces (o frecuencia) que crees que ocurra cada valor:
<table>
<tr><td>Caras</td><td></td><td></td><td></td><td></td><td>Total</td></tr>
<tr><td>Frecuencia</td><td></td><td></td><td></td><td></td><td></td></tr>
</table>
B) Anota la probabilidad que asignas a la ocurrencia de cada valor de la variable:
<table>
<tr><td>Caras</td><td>0</td><td>1</td><td>2</td><td>3</td><td>Total</td></tr>
<tr><td>Probabilidad</td><td></td><td></td><td></td><td></td><td></td></tr>
</table></td><td>2.1 Determinar las probabilidades asociada a los valores de una variable aleatoria discreta desde un enfoque frecuencial (apartado A)

2.4 Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta (apartado B)</td></tr>
<tr><td>2.2 (Salazar, 2014) En un programa de televisión los concursantes hacen girar una flecha en una ruleta como en la figura y cada vez que</td><td>2.2 Determinar las probabilidades asociada a los valores de una</td></tr>
</table>

<table>
<tr>
<td>la flecha se detiene en la zona de color rojo, se gana un premio y el concurso se trata de hacer girar la flecha 3 veces. En base a esta situación responde las siguientes preguntas:

A) ¿Cuál es la probabilidad de que un participante gane dos premios? Y ¿Cuál es la probabilidad de que un concursante no gane ningún premio?
B) Escribe en la tabla siguiente la función de probabilidad de la variable aleatoria que relaciona el espacio muestral con el número de premios que gana un concursante:
<table>
<tr><td>X</td><td>0</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>f(x)</td><td></td><td></td><td></td><td></td></tr>
</table>
C) En un gráfico represente la función de probabilidad anterior.</td>
<td>variable aleatoria discreta desde un enfoque clásico
(apartado A)

2.3 Definir la función de probabilidad de una variable aleatoria, utilizando la terminología P (X = xi).
(apartado B)

2.5 Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta (apartado C)</td>
</tr>
</table>

Nota. Elaboración propia.

Para representar la situación-problema 2.2 los expertos eligieron el ítem 2.2A. En él no es inmediata la obtención de las probabilidades solicitadas, sino que es necesario razonar e interpretar el enunciado (no ganar ningún premio quiere decir ganar cero premios). También permite aplicar diferentes elementos asociados al tópico de probabilidad como los sucesos aleatorios independientes, la combinatoria, la regla del producto, entre otros.

Sobre la situación-problema 2.3, los participantes escogieron para caracterizarla el ítem 2.2B, donde se propone un tipo de lenguaje específico para representar la función de probabilidad (lenguaje tabular). Por lo general dicho lenguaje de representación es más cercano a estudiantes que el simbólico, además facilita la respuesta dado que se entrega el dominio de la función requerida.

En relación con la situación-problema 2.4, los expertos prefirieron el ítem 2.1B. Esto se debió a que su enunciado proporciona facilidades, tanto en entregar una tabla donde se espera que sea representada la función de probabilidad, como en proveer completada la fila superior de esta tabla, correspondiente a los elementos del dominio de dicha función. Aún más, la tabla contiene una columna que hace referencia al total de las probabilidades, la cual sirve para corroborar si la función identificada, cumple con una de las condiciones para ser de probabilidad (suma de sus probabilidades igual a uno). De esta manera los aspectos mencionados favorecen a comprender lo solicitado y responder de manera correcta.

Finalmente, los expertos para representar la situación-problema 2.5, optaron por el ítem 2.2C. Esto se debe a su vinculación con el contexto de ítem 2.2B (situación-problema 2.3), el cual proporciona todos los elementos del dominio de la función de probabilidad, pero haciendo referencia explícita a los valores de X (variable aleatoria discreta). Esta información sin duda favorece a la resolución de la situación-problema 2.5, específicamente ayuda a asociar el eje horizontal del gráfico con el dominio de dicha función (valores de la variable aleatoria) y su eje vertical con el recorrido (probabilidades).

4.2.4.3 Estimación de la validez de contenido por juicio de expertos

A cada uno de los ítems seleccionados a través de juicio de expertos, se le calculó el CVC de Hernández-Nieto (2002) y los resultados del procedimiento se exponen en la Tabla 4.8. Así es posible afirmar que los ítems elegidos pueden ser conservados debido a que cada uno poseen un coeficiente superior a 0,70, es decir, valora adecuadamente la situación-problema respectiva. También el conjunto final de ítems obtuvo en promedio un CVC igual a 0,8, este valor indica que aquellos ítems evalúan idóneamente la comprensión de variable aleatoria y sus aplicaciones en las distribuciones de probabilidad binomial y normal en el contexto escolar chileno. En consecuencia, se puede sugerir la utilización de la GSP-VADP como un instrumento de evaluación en una investigación en educación o para estimar el avance en la comprensión de los estudiantes en los temas señalados.

Tabla 4.8

Valor de CVC de cada ítem que integra la versión final de la GVC-VADP

Situación-Problema	Ítem	CVC	Situación-Problema	Ítem	CVC
1. 1	1	0,87	5.3	15B	0,9
1. 2	1	0,87	5.4	14.2A	0,87
1.3	2B	0,83	5.5	15 C	0,8
1.4	3.2C	0,8	5.6	17	0,67
2.1	5.1A	0,77	6.1	20	0,83
2.2	6.1A	0,8	6.2	20	0,83
2.3	6.1B	0,83	6.3	20E y F	0,92
2.4	5.1B	0,83	6.4	20A, B, C y D	0,9
2.5	6.1C	0,87	7.1	23	0,83
3.1	10A	0,77	7.2	26	0,87
3.2	11B	0,77	7.3	29	0,9
3.3	11B	0,83	7.4	32	0,77
4.1	8A	0,8	8.1	35.A	0,8
4.2	8A	0,94	8.2	34.B	0,83
4.3	8B	0,9	6.1	37	0,9
5.1	13.1B	0,87	7.1	40	0,93
5.2	15C	0,8	Media		0,87

Nota. Elaboración propia.

4.2.5 Conclusiones

La presente investigación permitió crear una guía integrada por diversas tareas, actividades o ejercicios para comprender la variable aleatoria y distribuciones de probabilidad binomial y normal, según el currículo escolar chileno, denominada GSP-VADP. Los resultados de este estudio muestran que la guía está constituida por 33 situaciones-problemas sobre dichos temas, identificadas a partir de un análisis riguroso a documentos curriculares vigente, principalmente de Chile (MINEDUC, 2015; MINEDUC, 2016; MINEDUC, 2019a; MINEDUC, 2019b; MINEDUC, 2019c), además de Estados Unidos (NGACBP y CCSSO, 2010; Franklin et al., 2005; NCTM, 2000). También esta herramienta presenta importantes avances en torno a la enseñanza de los tópicos probabilísticos señalados:

- Amplía la diversidad de situaciones-problemas sobre variable aleatoria identificadas por Ortiz (2002), siendo estas representativas de la educación escolar actual.
- Propone situaciones-problemas en torno a la distribución binomial, fundamentales para la educación escolar e inicio de la universitaria, que no han sido explicitadas como tal objeto matemático en estudios previos.
- Amplía las diferentes situaciones-problemas relativas a la distribución normal reconocidas por Valverde (2017), coherentes con la educación probabilística fomentada en documentos curriculares vigentes.

Por tanto, la Guía de Situaciones-Problemas sobre Variable Aleatoria y sus Aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar, posee información valiosa que puede servir a los profesores de matemáticas que ejercen su labor en las escuelas, para orientar su trabajo y diseñar propuestas de enseñanzas en torno a los temas en cuestión.

Por otra parte, los resultados del presente estudio indican que es viable utilizar la GSPVADP para identificar ítems que representen las situaciones-problemas que la constituyen, por medio de expertos. Debido a que el conjunto final de ítems, que ha sido seleccionados por los seis participantes, obtuvo en promedio que es bueno con un coeficiente de validez y concordancia superior a 0,8. También este conjunto de ítems complementa las actividades o tareas propuestas para la enseñanza a nivel escolar de la variable aleatoria (Bizet y Ramos, 2022; Doukhan y Gueudet, 2019), distribución binomial (Alvarado y Retamal, 2014) y distribución normal (Salinas et al., 2018), ya que de manera articulada integra situaciones-problemas sobre variable aleatoria y su aplicación en distribuciones de probabilidad.

Finalmente, esta investigación se proyecta en llevar a cabo un análisis a priori de las situaciones-problemas que componen el conjunto final de ítems. De esta manera, establecer de forma explícita el lenguaje, conceptos, propiedades, procedimientos y argumentos puestos en juego en su resolución, los cuales caracterizan el significado institucional de la variable aleatoria y aplicaciones en distribuciones de probabilidad. Posteriormente, se espera aplicar aquellos ítems a estudiantes que cursan último grado de educación escolar (17 a 19 años) o comienzan su formación universitaria, con el propósito de centrarse en identificar el significado personal (logrado) de los tópicos en cuestión.

Referencias

Alvarado, H. y Batanero, C. (2007). Dificultades de Comprensión de la Aproximación Normal a la Distribución Binomial. *Números, revista de didáctica de las matemáticas, 67*, 1-7.

Alvarado, H. y Retamal, M. (2014). Representaciones de la distribución de probabilidad binomial. En L. Andrade (Ed.), *Memorias del I Encuentro Colombiano de Educación Estocástica* (pp. 98-109). Asociación Colombiana de Educación Estocástica.

Amrani, H. y Zaki, M. (2015). Student's conceptual difficulties with respect to the notion of random variable. *International Journal of Education, Learning and Development, 3*(9), 82-98.

Barraza, M. (2007). La consulta a expertos como estrategia para la recolección de evidencias de validez basado en el contenido. *Investigación Educativa Duranguense, 2*(7), 5-14.

Batanero, C. (2004) Ideas estocásticas fundamentales ¿qué contenido se debe enseñar en la clase de probabilidades? En J. Fernández, M. Sousa y S. Ribeiro (Eds.), *Actas do I Encontro de Probabilidades e Estatística na Escola* (pp. 9-30). Centro de Investigação em Educação da Universidade do Minho.

Bizet, V. y Ramos, E. (2022). Valoración de una situación didáctica para la enseñanza de variable aleatoria y distribución de probabilidad en la educación secundaria chilena. *Innovaciones Educativas, 24*(36), 21-36.

Chacón, A., García, G., Rupín, P., Setz, J. y Villena, M. (2018). *Texto del estudiante de matemática 2° medio*. Ediciones SM.

Chevallard, Y. (1980). The didactics of mathematics: its problematic and related research. *Recherches en Didactiques des Mathématiques, 1*, 146-157.

Chevallard, Y. (1991). *La transposición didáctica: del saber sabio al saber enseñado*: Aique.

Colomé, S. (28 de diciembre de 2015). *¿Cuál es la probabilidad real del empate de la CUP?*. La Vanguardia. https://www.lavanguardia.com/politica/20151228/301071423295/probabilidad-empate-cup.html

Dinges, H. (2005). Variables, in particular random variables. En M. Hoffmann, J. Lenhard y F. Seeger (Eds.), *Activity and sign grounding mathematics education* (pp. 305–311). Springer.

Doukhan, C. y Gueudet, G. (2019). Students' difficulties at the secondary-tertiary transition: the case of random variables. En U. Jankvist, M. Van Den Heuvel-Panhuizen y M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 2464-2471). Freudenthal Group y Freudenthal Institute of the Utrecht University y ERME.

Fernández, F., Andrade, L. y Sarmiento, B. (2013). Rehaciendo el camino hacia la comprensión de la variable aleatoria. En P. Perry, C. Samper, Ó. Molina, L. Camargo, A. Echeverry, F. Fernández y B. Sarmiento (Eds.), *Aportes investigativos para el diseño curricular en geometría y estadística* (pp. 93–169). Universidad Pedagógica Nacional.

Flores, B., García, J. y Sánchez, E. (2014). Avances en la calidad de las respuestas a preguntas de probabilidad después de una actividad de aprendizaje con tecnología. En M. González, M. Codes, D. Arnau y T. Ortega (Eds.), *Investigación en Educación Matemática XVIII* (pp. 307-316). SEIEM.

Font, V., Godino, J. y Gallardo, J. (2013). The emergence of objects from mathematical practices. *Educational Studies in Mathematics, 82*(1), 97-124.

Franklin, C., Kader, G., Newborn, D., Moreno, J., Peck, R., Perry, M. y Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-k–12 curriculum framework*. American Statistical Association.

Gal, I. (2002). Adult's statistical literacy: meanings, components, responsibilities. *International Statistical Review, 70*(1), 1– 25.

García-García, J., Fernández, N., Arredondo, E. e Imilpán, I. (2022). The Binomial Distribution: Historical Origin and Evolution of Its Problem Situations. *Mathematics, 10*, 1-28.

Godino, J. (1999). Implicaciones metodológicas de un enfoque semiótico- antropológico para la investigación en didáctica de las matemáticas. En T. Ortega (Ed.), *Investigación en Educación Matemática III* (pp. 196-212): SEIEM.

Godino, J. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, *22*(2.3), 237-284.

Godino, J. y Batanero, C (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, *14*(3), 325-355.

Godino, J., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education*, *39*(1-2), 127-135.

González, Y. y Ojeda, A. (2017). Comprensión de la distribución normal en bachillerato. En L. Serna (Ed.), *Acta Latinoamericana de Matemática Educativa, 30* (pp. 207-2017). CLAME.

Guerrero, H., Batanero, C. y J. M. Contreras. (2016). Conocimientos sobre esperanza matemática en alumnos de bachillerato. En F. España (Ed.), *Actas del XVI congreso de enseñanza y aprendizaje de las matemáticas* (pp. 26-35). Sociedad Andaluza de Educación Matemática THALES.

Heitele, D. (1975). An epistemological view of fundamental stochastic ideas. *Educational Studies in Mathematics*, *6*(2), 187–205, 1975.

Hernández-Nieto, R. (2002). *Contributions to Statistical Analysis.* Universidad de los Andes.

Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la investigación.* (6ª ed.). Editorial McGraw Hill Education.

Krippendorff, K. (1990). *Metodología de análisis de contenido: teoría y práctica*. Paidós.

Ministerio de Educación de Chile. (2009). *Currículum: Objetivos fundamentales y contenidos mínimos obligatorios de la educación básica y media*. Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2015). *Bases Curriculares 7° básico a 2° medio. Gobierno de Chile.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2016). *Matemática, programa de estudio segundo medio.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019a). Chile. *Bases Curriculares 3º y 4º medio.* Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019b). *Programa de estudio matemática 3° medio para formación general*. Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019c). *Programa de estudio matemática 4° medio para formación general.* Unidad de Currículum y Evaluación.

Morales, F., Vega, M., Aguilar, M. Gúmera, C. y Marchant, P. (2008). *Texto manual matemática 3º y 4º medio – proyecto explor@ndo*. Ediciones SM.

Muñoz, G., Gutiérrez, V. y Muñoz, S. (2013). T*exto matemática cuarto año medio.* Santillana.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.
National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standars for Mathematics.* NGACBP and CCSSO.
Norambuena, P., Osorio, G., Romante, M. y Díaz, J. (2019). *Cuaderno de actividades de matemática 3° y 4° medio*. Ediciones SM.
Ortiz, J. (2002). *La probabilidad en los libros de texto.* Grupo de Investigación en Educación Estadística de la Universidad de Granada.
Osorio, G., Norambuena, P., Romante, M., Gaete, D., Díaz, J., Celedón, J., Morales, K., Ortiz, N., Ramírez, P., Barrera, R. y Hurtado, Y. (2019). *Texto del estudiante de matemática 3° y 4° medio*. Ediciones SM.
Ramírez, N. (28 de diciembre de 2018). *Siempre el 16% obtiene más de 600 puntos en la PSU: Demre explica cómo funciona la prueba. Emol*. https://www.emol.com/noticias/Nacional/2018/12/28/932495/Siempre-el-16-obtiene-mas-de-600-puntos-en-la-PSU-DEMRE-explica-como-funciona-la-prueba.html
Rivas, H. (2014). *Idoneidad didáctica de procesos de formación estadística de profesores de educación primaria* [Tesis de doctorado no publicada]. Universidad de Granada.
Ruiz, B. (2006). *Un acercamiento cognitivo y epistemológico a la didáctica del concepto de variable aleatoria* [Tesis de maestría no publicada]. Instituto Politécnico Nacional.
Ruiz, B. (2013). *Análisis epistemológico de la variable aleatoria y comprensión de objetos matemáticos relacionados por estudiantes universitarios* [Tesis de doctorado no publicada]. Universidad de Granada.
Salazar, R. (2014). *La variable aleatoria con probabilidad desde la perspectiva de la teoría APOE* [Tesis de maestria no publicada]. Pontificia Universidad Católica de Valparaíso.
Salinas, J., Valdez, J. y Salinas-Hernández, U. (2018). Un acercamiento a la metodología lesson study para la enseñanza de la distribución normal. En L. Rodríguez-Muñiz, L. Muñiz-Rodríguez, A. Aguilar-González, P. Alonso, F. García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 525- 534). SEIEM.
Sánchez, E. y Carrasco, G. (2018). El razonamiento probabilístico de estudiantes de bachillerato en actividades de distribución binomial. En L. Rodríguez-Múniz, L. Múñiz-Rodríguez, Á. Aguilar-González, P. Alonso, F. García y A. Bruno (Eds.*), Investigación en Educación Matemática XXII* (pp. 535- 543). SEIEM.
Sánchez, E. y Landín, P. (2014). Levels of probabilistic reasoning of high school students about binomial problems. En E. Chernoff y B. Sriraman (Eds.), *Probabilistic thinking presenting plural perspectives* **(**pp. 581-597**)**. Springer.
Stahl, S. (2006). *The Evolution of the Normal Distribution. Mathematics Magazine, 79*(2), 96-113.
Tauber, L. (2001). *La construcción del significado de la distribución normal a partir de actividades de análisis de datos* [Tesis de doctorado no publicada]. Universidad de Sevilla.
Valverde, M. (2017). *Un estudio de la presentación de la distribución normal en los textos de bachillerato* [Tesis de maestría no publicada]. Universidad de Granada.
Vilca, M. (2015). *Tipificación de los errores que se presentan al identificar una variable aleatoria de distribución binomial en problemas contextualizados* [Tesis de maestría no publicada]. Pontificia Universidad Católica del Perú.

4.3 Análisis a priori de la versión inicial del instrumento

La versión inicial del cuestionario está organizado en ocho campos de problemas (C-P), cada uno compuesto por dos a seis situaciones-problemas (S-P), donde un ítem (I) valora a lo más tres de estas tareas. En cada uno de los ítems se han descrito las prácticas matemáticas que permiten obtener su solución correcta e identificado los objetos matemáticos primarios involucrados en aquellas prácticas.

4.3.1 C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico

El C-P_1 está integrado por cuatro situaciones-problemas (ver Figura 3): el $I_{1.1}$ valora la S-$P_{1.1}$ diferenciar entre variables aleatorias y variables con dependencia funcional y S-$P_{1.2}$ definir variables aleatorias finitas en experimentos aleatorios (Bizet, et al. 2023c, p.187); el $I_{1.2}$ representa la S-$P_{1.3}$ identificar dominio de una variable aleatoria finita (Bizet, et al. 2023c, p.187); y el $I_{1.3}$ evalúa la S-$P_{1.4}$ identificar recorrido o valores de una variable aleatoria finita (Bizet, et al. 2023c, p.187).

Figura 3

Situaciones-problemas que componen el C-P_1

1. De las situaciones que se describen a continuación: a) Lanzar dos dados para determinar el valor absoluto de la diferencia de los resultados. b) Tomar un artículo de un lote de 20 artículos de una fábrica, para determinar si está o no defectuoso. c) En la caída libre de un cuerpo, la relación entre el espacio recorrido y el tiempo desde su lanzamiento (donde la gravedad es $9{,}8\frac{\text{metros}}{\text{segundos}^2}$) d) Lanzar una moneda 4 veces y anotar el resultado que más se repite entre cara y sello. 1.1 Establece cuáles pueden ser variables aleatorias y descríbalas en palabras. 1.2 Enlista todos los posibles resultados de cada experimento asociado a las variables aleatorias descritas anteriormente. 1.3 Enlista los valores de cada variable aleatoria descrita, además diseña un diagrama que relacione todos los posibles resultados de cada experimento aleatorio (ítem 1.2) con estos valores.

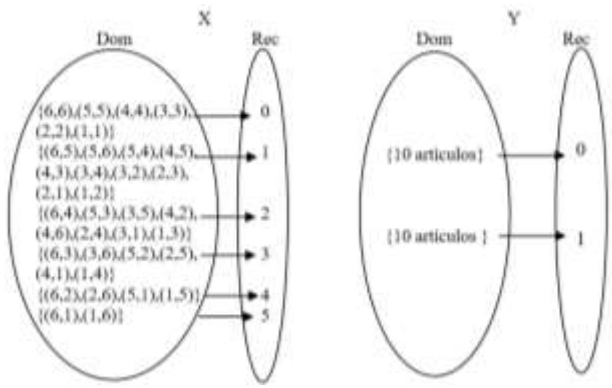
Nota. Adaptado de Fernández et al. (2013), Flores et al. (2014) y Salazar (2014), y extraído de Flores et al. (2014, p.311).

Para resolver el I_1, el estudiante debe reconocer que solo las situaciones (a) y (b) permiten definir una variable aleatoria. Luego para (a) definir la variable aleatoria X: valor absoluto de la diferencia entre los números obtenidos al lanzar dos dados, determinar su Dom(X)={(6,6), (5,5), (4,4), (3,3), (2,2), (1,1), (6,5), (5,6), (5,4), (4,5) (4,3), (3,4), (3,2), (2,3), (2,1), (1,2), (6,4), (5,3), (3,5), (4,2), (4,2), (2,4), (3,1), (1,3), (6,3), (3,6), (5,2), (2,5), (4,1), (1,4), (6,2), (2,6), (5,1), (1,5), (6,1), (1,6)}, su Rec(X)= {0, 1, 2, 3, 4, 5} y su diagrama sagital (ver Figura 4).

Análogamente para (b), definir la variable aleatoria Y que representa el estado del artículo en un determinado momento, siendo 1 si el artículo está defectuoso y 0 en caso contrario, determinar su Dom(Y)= {artículo 1, artículo 2,…., artículo 20} y su Rec(Y)={0, 1} donde 1 significa defectuoso y 0 no defectuoso. Un ejemplo de diagrama se expone en la Figura 4, el cual representa que 10 artículos están defectuoso y los otros 10 no.

Figura 4

Diagrama sagital de la variable aleatoria X e Y



Nota. Elaboración propia.

Es importante señalar que la situación que se describe en (c) es un fenómeno determinista que permite definir variables con dependencia funcional, donde la variable independiente X es el tiempo desde el lanzamiento del cuerpo y la variable dependiente Y es el espacio recorrido, siendo la gravedad g una constante igual a 9,8 $\frac{\text{metros}}{\text{segundos}^2}$: $Y = \frac{1}{2} \cdot g \cdot X^2$. Además la situación descrita en (d) no permite definir una variable aleatoria, aunque lanzar una moneda es un fenómeno o experimento aleatorio repetido 4 veces que permitiría definir algunas variables aleatorias como: i) número de caras obtenidas; y ii) número de sellos obtenidos.

Los diversos objetos matemáticos involucrados en las anteriores prácticas desarrolladas para resolver el $I_{1.1}$, $I_{1.2}$ e $I_{1.3}$ son expuestos en la Tabla 9.

Tabla 9

Objetos matemáticos involucrados en la solución del $I_{1.1}$, $I_{1.2}$ e $I_{1.3}$

Tipo de objeto matemático	Objeto matemático	Ítem 1.1	Ítem 1.2	Ítem 1.3
Situaciones-problemas	Diferenciar entre variables aleatorias y variables con dependencia funcional	X		
	Definir variables aleatorias finitas en experimentos aleatorios	X		
	Identificar dominio de una variable aleatoria finita		X	
	Identificar recorrido o valores de una variable aleatoria finita			X
Lenguaje	Verbal	X	X	X

	Simbólico		X	X
	Gráfico			X
Conceptos	Experimento aleatorio	X	X	X
	Suceso aleatorio	X	X	X
	Espacio muestral	X	X	X
	Función (dominio y recorrido)	X	X	X
	Conjunto finito	X		X
	Variable aleatoria	X	X	X
	Variable dependiente e independiente	X		
Proposiciones	Caracterización de la variable aleatoria y variables con dependencia funcional	X	X	
	Caracterización de variables aleatorias discreta y continua			X
Procedimientos	Cálculo de cardinalidad del espacio muestral		X	
	Construcción del espacio muestral		X	X
	Partición disjunta del espacio muestral			X
	Construcción de un diagrama sagital			X
Argumentos	Verbal-deductivo	X	X	
	Mediante representación gráfica			X

Nota. Elaboración propia.

4.3.2 C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria

El C-P_2 está constituido por cinco situaciones-problemas (ver Figuras 5 y 6): el $I_{2.1a}$ e $I_{2.1b}$ valoran respectivamente la S-$P_{2.1}$ determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial y S-$P_{2.4}$ representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta (Bizet, et al. 2023c, p.187); el $I_{2.2a}$ representa la S-$P_{2.2}$ determinar las probabilidades asociada a los valores de una variable aleatoria discreta desde un enfoque clásico (Bizet, et al. 2023c, p.187); el $I_{2.2b}$ representa la S-$P_{2.3}$ definir la función de probabilidad de una variable aleatoria utilizando la terminología P $(X = x_i)$ (Bizet, et al. 2023c, p.187); y el $I_{2.2c}$ evalúa la S-$P_{2.5}$ representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta (Bizet, et al. 2023c, p.187).

Figura 5

S-$P_{2.1}$ y S-$P_{2.4}$ que componen el C-P_2

2.1 Llamaremos lanzamiento a la acción de lanzar tres monedas al aire al mismo tiempo (de preferencia de la misma denominación). Ahora imagínate que se realizan 1000 lanzamientos y en cada uno de ellos se observa la variable "número de caras que ocurren".

a) En la siguiente tabla anota en la fila superior, los posibles valores de la variable y en la fila inferior, el número de veces (o frecuencia) que crees que ocurra cada valor:

Caras					Total
Frecuencia					

b) Anota la probabilidad que asignas a la ocurrencia de cada valor de la variable:

Caras	0	1	2	3	Total
Probabilidad					

Nota. Extraída de Flores et al. (2014, p.311).

Sobre la resolución del $I_{2.1}$, en el apartado a) el estudiante requiere reconocer el recorrido de la variable X: número de caras obtenidas, es decir, Rec(X)= {0,1,2,3}, luego determinar su distribución de frecuencias (0,125 o 125; 0,375 o 375; 0,375 o 375 y 0,125 o 125 correspondientemente), que varía levemente respecto a los valores teóricos de su distribución de probabilidad, y establecer la suma de las frecuencias (1 o 1000), para completar la tabla presentada. En el apartado b) el estudiante solo necesita determinar la probabilidad teórica de cada valor de la variable aleatoria $P(0) = \frac{125}{1000} = 0{,}125$ o 12,5%; $P(1) = \frac{375}{1000} = 0{,}375$ o 37,5%; $P(2) = \frac{375}{1000}$ y $P(3) = \frac{125}{1000}$ cuya suma es 1(o 100%), así completar la tabla de probabilidad propuesta.

Figura 6

S-$P_{2.2}$, S-$P_{2.3}$ y S-$P_{2.5}$ que componen el C-P_2

2.2 En un programa de televisión los concursantes hacen girar una flecha en una ruleta como en la figura y cada vez que la flecha se detiene en la zona de color rojo, se gana un premio y el concurso se trata de hacer girar la flecha 3 veces. En base a esta situación responde las siguientes preguntas:

a) ¿Cuál es la probabilidad de que un participante gane dos premios? Y ¿Cuál es la probabilidad de que un concursante no gane algún premio?

b) Escribe en la tabla siguiente la función de probabilidad de la variable aleatoria que relaciona el espacio muestral con el número de premios que gana un concursante.

X	0	1	2	3
f(x)				

c) En un gráfico represente la función de probabilidad anterior.

Nota. Extraída de Salazar (2014, p.56-57) y elaboración propia.

Para resolver el $I_{2.2}$, el estudiante debe establecer la función de probabilidad, entonces puede realizar la estrategia de utilizar los conceptos de combinatoria y sucesos independientes y la propiedad de la regla del producto:

$$\text{P(ganar 0 premios al girar la flecha de la ruleta 3 veces)} = C_0^3 . \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} = 1 \cdot \frac{27}{64} = 0{,}4218 \text{ o } 42{,}18\%$$

$$\text{P(ganar 1 premio al girar la flecha de la ruleta 3 veces)} = C_1^3 \cdot \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} = 3 \cdot \frac{9}{64} = \frac{27}{64}$$

$$\text{P(ganar 2 premios al girar la flecha de la ruleta 3 veces)} = C_2^3 \cdot \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} = 3 \cdot \frac{3}{64} = \frac{9}{64} = 0{,}1406 \text{ o } 14{,}06\%$$

$$\text{P(3 premios al girar la flecha 3 veces)} = C_3^3 \cdot \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{64} = \frac{1}{64} = 0{,}0156 \text{ o } 1{,}56.$$

Otra estrategia que el estudiante puede desarrollar es definir la variable aleatoria X: número de premios que gana un concursante al girar la flecha de la ruleta tres veces, la cual se distribuye de forma binomial, y determinar su función de probabilidad:

$P(X = 0) = C_0^3 \cdot \left(\frac{1}{4}\right)^0 \cdot \left(\frac{3}{4}\right)^3 = 1 \cdot \frac{27}{64} = 0{,}4218$ o $42{,}18\%$

$P(X = 1) = C_1^3 \cdot \left(\frac{1}{4}\right)^1 \cdot \left(\frac{3}{4}\right)^2 = 3 \cdot \frac{9}{64} = \frac{27}{64}$

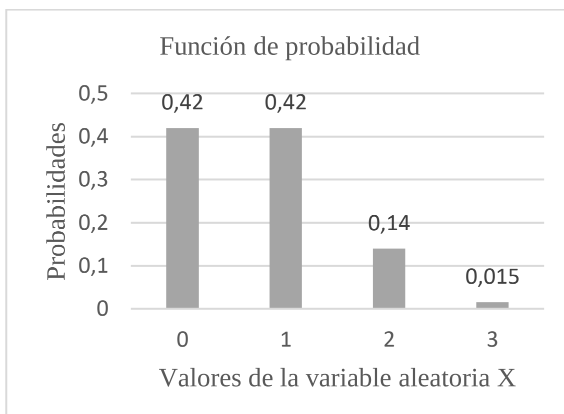
$P(X = 2) = C_2^3 \cdot \left(\frac{1}{4}\right)^2 \cdot \left(\frac{3}{4}\right)^1 = 3 \cdot \frac{3}{64} = \frac{9}{64} = 0{,}1406$ o $14{,}06\%$

$P(X = 3) = C_3^3 \cdot \left(\frac{1}{4}\right)^3 \cdot \left(\frac{3}{4}\right)^0 = 1 \cdot \frac{1}{64} = \frac{1}{64} = 0{,}0156$ o $1{,}56$.

Aquellos valores permiten que el estudiante conozca la probabilidad que un participante gane dos premios P(2)=0,1406 y no gane premio P(0)=0,4218, además de completar la tabla de probabilidad propuesta. Luego requiere construir un gráfico de barras donde se represente la función de probabilidad de X (ver Figura 7).

Figura 7

Función de probabilidad de X



Nota. Elaboración propia.

La Tabla 10 expone los objetos matemáticos que intervienen en las prácticas anteriormente descritas realizadas para resolver el $I_{2.1}$ e $I_{2.2}$.

Tabla 10

Objetos matemáticos involucrados en la solución del $I_{2.1}$ *e* $I_{2.2}$

Tipo de objeto matemático	Objeto matemático	Ítem 2.1a	2.1b	2.2a	2.2b	2.2c
Situaciones-problemas	Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial	X				
	Determinar las probabilidades asociada a los valores de			X		

	una variable aleatoria discreta desde un enfoque clásico					
	Definir la función de probabilidad de una variable aleatoria utilizando la terminología P (X = x_i)				X	
	Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta		X			
	Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta					X
Lenguaje	Verbal	X	X	X	X	X
	Numérico	X	X	X	X	
	Simbólico		X	X	X	X
	Gráfico					X
	Tabular	X	X		X	
Conceptos	Experimento aleatorio	X	X	X	X	X
	Suceso aleatorio	X	X	X	X	X
	Sucesos independientes			X	X	X
	Espacio muestral	X	X	X	X	X
	Variable aleatoria	X	X	X	X	X
	Conjunto finito	X	X	X	X	X
	Enfoque frecuencial de probabilidad	X				
	Enfoque clásico de probabilidad		X	X	X	X
	Distribución de frecuencias	X				
	Función de probabilidad		X	X	X	X
	Combinatoria			X	X	X
	Distribución binomial			X	X	X
	Función de probabilidad de la binomial			X	X	X
Proposiciones	Caracterización de variables aleatorias discreta y continua	X	X	X	X	X
	Convergencia, al crecer el número de ensayos la frecuencia relativa se va estabilizando	X				
	Regla de Laplace		X	X	X	X
	Caracterización de la variable aleatoria mediante su función de probabilidad		X	X	X	X
	Propiedades de la función de probabilidad		X	X	X	X
	Caracterización de la distribución binomial			X	X	X
	La probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento			X	X	X
	Regla del producto			X	X	X
Procedimientos	Partición disjunta del espacio muestral	X	X	X	X	X
	Reproducción de experimentos aleatorio conservando las condiciones iniciales	X				
	Cálculo de frecuencias (relativas o absolutas)	X				
	Cálculo de probabilidades con la regla de Laplace		X	X	X	X
	Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial			X	X	X
	Cálculo de probabilidad compuesta con la regla del producto			X	X	X
	Cálculo de combinaciones			X	X	X
	Construcción de un gráfico de barras					X
Argumentos	Verbal-deductivo	X	X	X	X	
	Mediante representación gráfica					X

Nota. Elaboración propia.

4.3.3 C-P_3 Reconocer la distribución de probabilidad de una variable aleatoria continua (función de densidad) como una herramienta que permite ver la variación aleatoria

El C-P_3 está compuesto por tres situaciones-problemas (ver Figura 8): el $I_{3.1}$ valora la S-$P_{3.1}$ representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua (Bizet, et al. 2023c, p.187); el $I_{3.2a}$ evalúa la S-$P_{3.3}$ determinar si la función dada es una función de probabilidad de una variable aleatoria continua(Bizet, et al. 2023c, p.187); y el $I_{3.2b}$ representa la S-$P_{3.2}$ calcular algunas probabilidades asociadas a una variable aleatoria continua (Bizet, et al. 2023c, p.187).

Figura 8

Situaciones-problemas que componen el C-P_3

3.1 Considera la función de densidad f(x) = 0,2 definida en el intervalo [0, 5] y construye la gráfica de f.
3.2 A partir de la función f, definida en el intervalo [–0,5; 1]

$$f(x) = \begin{cases} x+1 & -0,5 \leq x \leq 0 \\ -x+1 & 0 < x \leq 0,5 \\ 0,5 & 0,5 < x \leq 1 \end{cases}$$

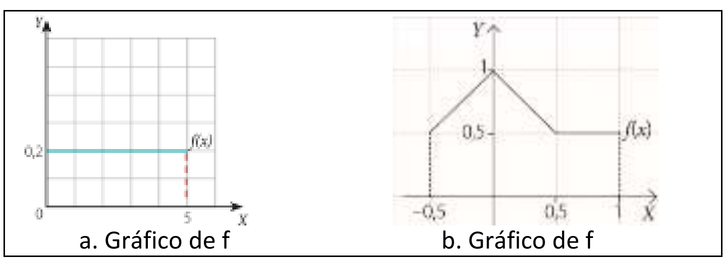
a) Determina si f puede ser la función de densidad de una variable aleatoria continua.
b) Calcula P (X = 0,5), P (X < 0), P (0,5 < X < 1) y P (X > 2).

Nota. Extraída de Muñoz et al. (2013, p.334) y adaptado de Muñoz et al. (2013).

Respecto a la resolución del $I_{3.1}$, el estudiante debe graficar en el plano cartesiano la función constante propuesta como se muestra en la Figura 9(a).

Figura 9

Gráficos de funciones del $I_{3.1}$ e $I_{3.2}$



Nota. Elaboración propia.

La resolución del $I_{3.2}$ involucra calcular integrales pero este concepto está excluido del currículo escolar chileno, entonces se sugiere trabajar con la interpretación geométrica de la integral definida. Por lo que el estudiante requiere graficar en el plano cartesiano la función por parte propuesta, como se expone en la Figura 9(b). Puntualmente en el apartado a), este debe comprobar si aquella función f satisface dos condiciones para ser función de densidad: 1. $f(x) \geq 0$, la cual se cumple ya que la gráfica de f siempre se encuentra sobre el eje X, en otras palabras, los valores que toma f(x) siempre son positivos; 2. el área bajo la curva de f es igual a 1, se comprueba su cumplimiento calculando el área bajo la curva, a partir del área de un rectángulo ($Á_{\text{rectángulo}} = 1{,}5 \cdot 0{,}5 = 0{,}75\text{u}^2$) y un triángulo ($Á_{\text{triángulo}} = \frac{1 \cdot 0{,}5}{2} = 0{,}25\ \text{u}^2$), así el $Á_{\text{total}} = 0{,}75 + 0{,}25 = 1\ \mu^2$. Por tanto, como se cumplen ambas condiciones, es posible afirmar que f puede ser la función de densidad de una variable aleatoria continua.

Para responder el apartado b) el estudiante necesita calcular algunas áreas de figuras geométricas, es decir, $\text{P}(\text{X} < 0) = Á_{\text{cuadrado}} + Á_{\text{triángulo}} = 0,25 + 0{,}125 = 0{,}375$ y $\text{P}(0{,}5 < \text{X} < 1) = Á_{\text{cuadrado}} = 0{,}25$, además de aplicar propiedades de la función de densidad: $\text{P}(\text{X} = 0{,}5) = 0$ debido a que la probabilidad de que una variable aleatoria continua tome un valor fijo siempre es 0; $P(\text{X} > 2) = 0$ dado que la función está definida en el intervalo [–0,5; 1] entonces cualquier probabilidad fuera de ese intervalo es 0.

Los diferentes objetos matemáticos movilizados en las anteriores prácticas llevadas a cabo para resolver el $I_{3.1}$ e $I_{3.2}$ son presentados en la Tabla 11.

Tabla 11

Objetos matemáticos involucrados en la solución del $I_{3.1}$ e $I_{3.2}$

Tipo de objeto matemático	Objeto matemático	Ítem $I_{3.1}$	$I_{3.2a}$	$I_{3.2b}$
Situaciones-problemas	Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua	X		
	Calcular algunas probabilidades asociadas a una variable aleatoria continua			X
	Determinar si la función dada es una función de probabilidad de una variable aleatoria continua		X	
Lenguaje	Verbal	X	X	X
	Numérico		X	X
	Simbólico	X	X	X
	Gráfico	X	X	X
Conceptos	Variable aleatoria	X	X	X
	Función (dominio y recorrido)	X	X	X
	Intervalo	X	X	X
	Función de densidad	X	X	X
	Área		X	X
Proposiciones	Caracterización de variables aleatorias discreta y continua	X	X	X
	Propiedades de la función de densidad	X	X	X
Procedimientos	Construcción de la gráfica de una función real en el plano cartesiano	X	X	X
	Cálculo de probabilidad de una variable aleatoria continua aplicando la función de densidad		X	X
	Cálculo de integral definida mediante su interpretación geométrica		X	X
Argumentos	Mediante representación gráfica	X	X	X

Nota. Elaboración propia.

4.3.4 C-P_4 Establecer la función de distribución de una variable aleatoria discreta

El C-P_4 está integrado por tres situaciones-problemas (ver Figura 10): el $I_{4.a}$ valora la S-$P_{4.1}$ determinar la probabilidad acumulada hasta algunos valores de la variable aleatoria discreta y S-$P_{4.2}$ definir la función de distribución de una variable aleatoria discreta(Bizet, et al. 2023c, p.187); y el $I_{4.b}$ representa la S-$P_{4.3}$ representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta(Bizet, et al. 2023c, p.187).

Figura 10

Situaciones-problemas que componen el C-P_4

4. Para el gráfico que muestra la función de probabilidad de una variable aleatoria X:

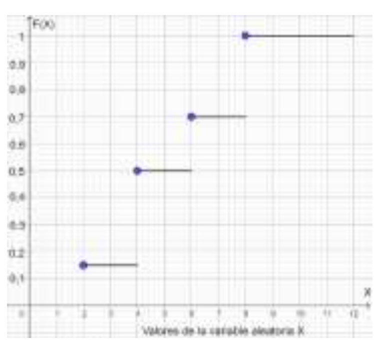
a) Escribe su función de distribución.
b) Construye un gráfico que represente su función de distribución.

Nota. Extraída de Chacón et al. (2018, p.292) y elaboración propia.

La resolución del I_4, requiere que el estudiante lea e interprete el gráfico de barras propuesto. A partir de la información extraída calcule las probabilidades acumuladas hasta cada valor de la variable aleatoria X y determine su función de distribución $P(X \leq 2) = F(2) = 0{,}15$; $P(X \leq 4) = F(4) = 0{,}5$; $P(X \leq 6) = F(6) = 0{,}7$ y $P(X \leq 8) = F(8) = 1$. Luego debe graficar aquella función en el plano cartesiano (ver Figura 11).

Figura 11

Función de distribución de la variable aleatoria X



Nota. Elaboración propia.

La Tabla 12 muestra los objetos matemáticos involucrados en las prácticas anteriormente descritas realizadas para resolver el I_4.

Tabla 12

Objetos matemáticos involucrados en la solución del I_4

Tipo de objeto matemático	Objeto matemático	Ítem 4.a	Ítem 4.b
Situaciones-problemas	Determinar la probabilidad acumulada de algunos valores de la variable aleatoria discreta	X	
	Definir la función de distribución de una variable aleatoria discreta	X	
	Representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta		X
Lenguaje	Verbal	X	X
	Numérico	X	
	Simbólico	X	X
	Gráfico	X	X
Conceptos	Variable aleatoria	X	X
	Función (dominio y recorrido)		X
	Función de probabilidad	X	X
	Función de distribución	X	X
Proposiciones	Caracterización de variables aleatorias discreta y continua	X	X
	Caracterización de la variable aleatoria mediante su función de probabilidad	X	X
	Propiedades de la función de probabilidad	X	X
	Caracterización de la variable aleatoria mediante su función de distribución	X	X
Procedimientos	Construcción de la gráfica de una función real en el plano cartesiano		X
	Lectura e interpretación de un gráfico de barras	X	X
	Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de distribución	X	X
Argumentos	Mediante representación gráfica	X	X

Nota. Elaboración propia.

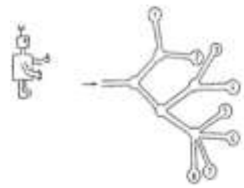
4.3.5 C-P_5 Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria

El C-P_5 está constituido por seis situaciones-problemas (ver Figuras 12, 13 y 14): el $I_{5.1}$ evalúa la S-$P_{5.1}$ calcular la media de una variable aleatoria discreta(Bizet, et al. 2023c, p.187); el $I_{5.2a}$ valora la S-$P_{5.4}$ calcular la desviación estándar de una variable aleatoria discreta(Bizet, et al. 2023c, p.187); el $I_{5.2b}$ representa la S-$P_{5.3}$ calcular la varianza de una variable aleatoria discreta (Bizet, et al. 2023c, p.187); el $I_{5.2c}$ valora la S-$P_{5.2}$ interpretar la media de una variable aleatoria discreta y S-$P_{5.5}$ interpretar la desviación estándar de una variable aleatoria discreta (Bizet, et al. 2023c, p.187); y el $I_{10.a}$ evalúa la S-$P_{5.6}$ identificar la media y desviación estándar de una variable aleatoria continua (Bizet, et al. 2023c, p.187).

Figura 12

S-P$_{5.1}$ que componen el C-P$_5$

5.1 Un robot es colocado ante un laberinto que empieza a explorar. En cada cruce el robot tiene tantas probabilidades de irse por un camino como por otro (Pero nunca vuelve por el camino que vino). Hay 8 metas al final de los 8 caminos (ver el dibujo)



Si ponemos 24 robots al principio del laberinto, ¿Cuántos piensas que acabarán en una de las metas 1 ó 2? ¿Por qué?

Nota. Extraída de Guerrero et al. (2016, p.2).

Para responder el I$_{5.1}$, el estudiante necesita definir la variable aleatoria X: número de robots (1 a 24) que pueden acabar en cada una de las metas 1 o 2 y calcular su media $E(X) = n \cdot P(1 \cup 2) = 24 \cdot \left(\frac{1}{4} + \frac{1}{4}\right) = \frac{24}{2} = 12$ obteniendo que 12 son los robots que terminarán en las metas señaladas.

Figura 13

S-P$_{5.2}$, S-P$_{5.3}$, S-P$_{5.4}$ y S-P$_{5.5}$ que componen el C-P$_5$

5.2 Se define la variable aleatoria X: número de mascotas que tiene un estudiante. La función probabilidad asociada es:

x_i	0	1	2	3
$P(x_i)$	0,325	0,375	0,225	0,075

a) ¿Cuál es la desviación estándar de la variable?

A) $\sqrt{1{,}55625}$ B) $\sqrt{0{,}9}$ *C)* $\sqrt{0{,}8475}$ D) $\sqrt{0{,}0525}$

b) ¿Cuál es la varianza de la variable?

c) Si en el curso 3 medio de un liceo, el número medio de mascotas que tiene un estudiante es 1 y la desviación estándar es 0,95 ¿Cómo interpretarías cada uno de los resultados?

Nota. Extraída de Osorio et al. (2019, p.166) y elaboración propia.

En el I$_{5.2}$ el estudiante requiere leer la tabla propuesta, calcular la varianza de la variable aleatoria X, esto es $\mathrm{Var}(X) = E(X^2) - \left(E(X)\right)^2 = 1{,}95 - 1{,}1025 = 0{,}8475$, y calcular la raíz cuadrada de este valor para obtener su desviación estándar $S(X) = \sqrt{0{,}8475} = 0{,}921$. Además, debe interpretar la media y desviación estándar de la variable aleatoria discreta propuestas (número de mascotas de un estudiante del curso 3 medio), es decir, que los estudiantes del curso 3 medio, en promedio, tienen una mascota y el número de mascotas de aquel curso se desvía, en promedio, 0,95 unidades respecto a la media 1.

Figura 14

S-P$_{5.6}$ que componen el C-P$_5$

10. Lee el siguiente titular, luego responda las preguntas propuestas

a) De manera resumida, explique cómo funciona la prueba PSU mencionando los datos representados en el gráfico

Nota. Elaboración propia y posee información de Ramírez (2018).

Para responder el I$_{10.a}$ el estudiante necesita identificar que la variable aleatoria continua puntaje obtenido por un individuo en la prueba PSU es normalizada, donde se utiliza una escala de puntaje estándar con media de 500 puntos y desviación estándar de 110 puntos, lo que implica que: i) entre los 390 y 610 puntos hay un determinado porcentaje de individuos, que es 68,26%; ii) entre 280 y 720 hay un 95,44%; y iii) entre 170 y 830 puntos se sitúa el 99,72% de la población.

La Tabla 13 expone los objetos matemáticos movilizados en las anteriores prácticas desarrolladas para resolver el I$_{5.1}$, I$_{5.2}$ e I$_{10.a.}$

Tabla 13

Objetos matemáticos involucrados en la solución del I$_{5.1}$, I$_{5.2}$ e I$_{10.a}$

Tipo de objeto matemático	Objeto matemático	Ítem 5.1	5.2a	5.2b	5.2c	10.a
Situaciones-problemas	Calcular la media de una variable aleatoria discreta	X				
	Interpretar la media de una variable aleatoria discreta				X	
	Calcular la varianza de una variable aleatoria discreta			X		
	Calcular la desviación estándar de una variable aleatoria discreta		X			
	Interpretar la desviación estándar de una variable aleatoria discreta				X	
	Identificar la media y desviación estándar de una variable aleatoria continua					X
Lenguaje	Verbal	X	X	X	X	X
	Numérico	X	X	X		
	Simbólico	X	X	X		X
	Gráfico	X				X
	Tabular		X	X		
Conceptos	Variable aleatoria	X	X	X	X	X
	Suceso aleatorio	X	X	X		

	Función de probabilidad	X	X	X		
	Intervalo					X
	Media	X			X	X
	Varianza		X	X		
	Desviación estándar		X		X	X
	Distribución normal					X
	Función de densidad de la normal					X
Proposiciones	Caracterización de variables aleatorias discreta y continua	X	X	X	X	X
	Caracterización de la variable aleatoria mediante su función de probabilidad	X	X	X		
	La media es un representante de los datos a partir de los que ha sido calculada y no tiene por qué ser igual a uno de los valores de los datos				X	
	Propiedad de los intervalos centrales (±3σ)					X
Procedimientos	Cálculo de probabilidades de una variable aleatoria discreta aplicando su función de probabilidad	X				
	Lectura de una tabla de probabilidad		X	X		
	Cálculo de valores de posición central y de dispersión asociados a una variable aleatoria discreta empleando sus expresiones algebraicas	X	X	X		
	Lectura e interpretación de una curva normal					X
Argumentos	Verbal-deductivo		X	X	X	
	Mediante representación gráfica	X				X

Nota. Elaboración propia.

4.3.6 C-P_6 La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real

El C-P_6 está compuesto por cuatro situaciones-problema (ver Figuras 15 y 16): el I_9 e I_6 evalúan la S-$P_{6.1}$ identificar situaciones que pueden modelarse a través de una distribución binomial (Bizet, et al. 2023c, p.187); también el I_6 valora la S-$P_{6.2}$ determinar los parámetros asociados a una distribución binomial (n, p y q) (Bizet, et al. 2023c, p.187); el I_{6a}, I_{6b}, I_{6c} e I_{6d} representan la S-$P_{6.4}$ determinar probabilidades asociadas a una distribución binomial de forma manual(Bizet, et al. 2023c, p.187); y el I_{6e} e I_{6f} evalúan la S-$P_{6.3}$ calcular la media y desviación estándar de una distribución binomial(Bizet, et al. 2023c, p.187).

Figura 15

S-$P_{6.1}$ que compone el C-P_6

9. Lea la siguiente noticia y responda a la pregunta propuesta

LA VANGUARDIA

Barcelona, 29/12/2015

¿Cuál es la probabilidad real del empate de la CUP?

Se ha generado una discusión en las redes sociales entre dos posibles ecuaciones que darían respuestas diferentes a esta incógnita

Las matemáticas están en cualquier aspecto de la vida, incluso en la política. O tal vez podría decirse que esta sea una de las disciplinas más marcadas por las leyes de los números. Y, de hecho, tras conocerse el empate en la asamblea de la CUP para investir a Artur Mas como president de la Generalitat, la pregunta resulta casi inevitable: ¿Cuál era la probabilidad de que esto ocurriera?

Como mínimo, esta es la cuestión que circula hoy por las redes sociales y algún matemático ya ha hecho los cálculos para dar la respuesta precisa. En la tercera y última votación, 3.030 militantes y simpatizantes cuperos se sorprendieron a sí mismos dando un resultado insólito, un sorprendente empate a 1.510 votos. Según el catedrático de Matemática Aplicada de la Universidad de Sevilla, Mario Bilbao, esta sería la ecuación: “3030/3031 = 0.99967, un suceso seguro al 99.967%”. Es decir, el empate es posible en un 0.00033014 de los casos.

Aunque resulte casi imposible, cabe decir que aún es más difícil que toque el Gordo de la Lotería de Navidad. En este caso, la probabilidad es de un 0,00001%, según explicó Adolfo Quirós, matemático y profesor de la Universidad Autónoma de Madrid (UAM), a La Vanguardia.com días antes del sorteo. Y la realidad es que en Roquetes del Mar hay muchas familias que han podido constatar que no es un mito recibir el premio más codiciado de la Lotería.

El matemático sevillano se ha valido de la Regla de Laplace para llegar a esta conclusión, desarrollada por Pierre-Simon Laplace en el siglo XVIII. Otros especialistas han llegado a esta misma conclusión en las redes sociales, como es el caso del sociólogo y doctor en Ciencias Económicas Salvador Cardús.

Pero tampoco en matemáticas hay nada exacto. En las redes circula otra ecuación para calcular esta probabilidad que da un resultado algo más elevado. Esta segunda tesis es la que defiende, por ejemplo, el usuario @shinnosuke, al que le da su apoyo, entre otros, el economista Xavier Sala-i-Martín. Según esta teoría, el empate es factible en un 1,45% de las veces

Sea cual sea la solución correcta, los hechos son los hechos, y en este caso han demostrado que la más remota probabilidad puede devenir la más sorprendente realidad.

A partir de la información entregada en la noticia ¿Cuál de los expertos podrían tener la respuesta correcta? Justifique su respuesta

Nota. Elaboración propia y posee información de Colomé (2015).

Para responder al I_9 el estudiante debe tener en cuenta que cuando se calculan probabilidades se consideran supuestos, en esta tarea existen dos: i) hay 3030 votos contabilizados, cada uno puede ser sí o no (0 o 1); ii) un empate se presenta cuando hay igual número de síes (1) que de noes (0). Luego este necesita reconocer en la situación propuesta las características de la distribución binomial, es decir, presenta un suceso dicotómico (votar sí o no) observado 3030 veces y los posibles resultados obtenidos en cada sufragio son independientes. Así requiere definir la variable aleatoria X: cantidad de síes en una muestra de tamaño 3030, donde $X \sim B\left(3030, \frac{1}{2}\right)$ e identificar que la probabilidad de obtener 1515 síes corresponde a la expresión $\mathrm{P}(1515) = \mathrm{C}_{1515}^{3030} \cdot \left(\frac{1}{2}\right)^{1515} \cdot \left(\frac{1}{2}\right)^{1515} = \frac{3030!}{(1515!)^2} \cdot \left(\frac{1}{2}\right)^{3030} = 0{,}0145 = 1{,}45\%$. De esta manera, determina que el usuario @shinnosuke y el economista Xavier Sala-i-Martín, proponen una respuesta correcta.

Figura 16

Situaciones-problemas que componen el C-P$_6$

6. Un profesor aplica a sus estudiantes un cuestionario sorpresa con cinco ítems de opción múltiple. Uno de los estudiantes no ha estudiado el material del cuestionario, y por tanto decide contestar los cinco ítems al azar, adivinando las respuestas sin leer las preguntas ni las opciones de respuestas. Hoja de respuestas del cuestionario. Instrucciones: encierre en un círculo la mejor respuesta de cada ítem 1. A B C 2. A B C 3. A B C 4. A B C 5. A B C Antes de ver las respuestas correctas del cuestionario y encontrar cómo le fue a este estudiante, se considerarán algunos hechos que podrían suceder si un cuestionario se contesta de esta forma. a) ¿Cuántos de los cinco ítems es probable contestar correctamente? b) Si el estudiante aprueba el cuestionario cuando responde al menos tres ítems correctamente ¿Cuál es la probabilidad de que apruebe? c) ¿Cuál es la probabilidad de elegir las respuestas acertadas de los cinco ítems? d) ¿Cuál es la probabilidad de elegir las respuestas incorrectas de los cinco ítems? e) Calcula la media de las respuestas correctas. f) Calcula la desviación estándar de las respuestas correctas.

Nota. Extraída de Alvarado y Retamal (2014, p.104), adaptada de Sánchez y Carrasco (2018) y elaboración propia.

En la resolución del I$_6$, el estudiante necesita identificar las características del contexto del problema, es decir, la situación expuesta presenta un suceso dicotómico (responder correcta o incorrectamente a un ítem) observado 5 veces y la respuesta a cada ítem es un suceso independiente. Después este debe definir la variable aleatoria X: número de respuestas correctas en el cuestionario y puede desarrollar la estrategia de aplicar el concepto de distribución binomial, identificando que $\mathrm{X} \sim \mathrm{B}\left(5, \frac{1}{3}\right)$, donde la probabilidad de elegir la respuesta correctamente de un ítem individual es $\mathrm{p} = \frac{1}{3}$, la probabilidad de seleccionar una respuesta incorrecta en un ítem es $\mathrm{q} = \frac{2}{3}$ y el número de ítems del cuestionario corresponde a $\mathrm{n} = 5$. Entonces el estudiante para calcular las probabilidades solicitadas emplea directamente la función de probabilidad de la binomial, pues el currículo chileno excluye trabajar con la tabla de la distribución binomial que facilita los valores de las probabilidades.

$\mathrm{P}(0) = \mathrm{C}_0^5 \cdot \left(\frac{1}{3}\right)^0 \cdot \left(\frac{2}{3}\right)^5 = \frac{32}{243} = 0{,}131$ o $13{,}1\%$

$\mathrm{P}(1) = \mathrm{C}_1^5 \cdot \left(\frac{1}{3}\right)^1 \cdot \left(\frac{2}{3}\right)^4 = \frac{80}{243} = 0{,}329 \text{ o } 32{,}9\%$

$\mathrm{P}(2) = \mathrm{C}_2^5 \cdot \left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^3 = \frac{80}{243} = 0{,}329 \text{ o } 32{,}9\%$

$\mathrm{P}(3) = \mathrm{C}_3^5 \cdot \left(\frac{1}{3}\right)^3 \cdot \left(\frac{2}{3}\right)^2 = \frac{40}{243} = 0{,}164 \text{ o } 16{,}4\%$

$\mathrm{P}(4) = \mathrm{C}_4^5 \cdot \left(\frac{1}{3}\right)^4 \cdot \left(\frac{2}{3}\right)^1 = \frac{10}{243} = 0{,}041 \text{ o } 4{,}1\%$

y $\mathrm{P}(5) = \mathrm{C}_5^5 \cdot \left(\frac{1}{3}\right)^5 \cdot \left(\frac{2}{3}\right)^0 = \frac{1}{243} = 0{,}004 \text{ o } 0{,}4\%$

a) Es más probable contestar una o dos respuestas correctas

b)$\mathrm{P}(\mathrm{X} \geq 3) = \mathrm{P}(3) + \mathrm{P}(4) + \mathrm{P}(5) = \mathrm{C}_3^5 \cdot \left(\frac{1}{3}\right)^3 \cdot \left(\frac{2}{3}\right)^2 + \mathrm{C}_4^5 \cdot \left(\frac{1}{3}\right)^4 \cdot \left(\frac{2}{3}\right)^1 + \mathrm{C}_5^5 \cdot \left(\frac{1}{3}\right)^5 \cdot \left(\frac{2}{3}\right)^0 = 10 \cdot \frac{4}{243} + 5 \cdot \frac{2}{243} + 1 \cdot \frac{1}{243} = \frac{51}{243} = 0{,}209 = 21\%$

c) $\mathrm{P}(5) = \mathrm{C}_5^5 \cdot \left(\frac{1}{3}\right)^5 \cdot \left(\frac{2}{3}\right)^0 = 1 \cdot \frac{1}{243} = 0{,}004 = 0{,}4\%$

d) $\mathrm{P}(0) = \mathrm{C}_0^5 \cdot \left(\frac{1}{3}\right)^0 \cdot \left(\frac{2}{3}\right)^5 = 1 \cdot \frac{32}{243} = 0{,}131 = 13{,}1\%$

Otra estrategia que el estudiante puede realizar es utilizar los conceptos de sucesos independientes y combinatoria y la propiedad de la regla del producto, así:

$\mathrm{P}(\text{elegir 0 respuestas correctas en 5 ítems}) = \mathrm{C}_0^5 \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{32}{243}$

$\mathrm{P}(\text{elegir 1 respuesta correcta en 5 ítems}) = \mathrm{C}_1^5 \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{80}{243}$

$\mathrm{P}(\text{elegir 2 respuestas correctas en 5 ítems}) = \mathrm{C}_2^5 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{80}{243}$

$\mathrm{P}(\text{elegir 3 respuestas correctas en 5 ítems}) = \mathrm{C}_3^5 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{40}{243}$

$\mathrm{P}(\text{elegir 4 respuestas correctas en 5 ítems}) = \mathrm{C}_4^5 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{10}{243}$

$\mathrm{P}(\text{elegir 5 respuestas correctas en 5 ítems}) = \mathrm{C}_5^5 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{243}$

a) Es más probable contestar una o dos respuestas correctas

b) $\mathrm{P}(\text{tres o más respuestas correctas en 5 ítems}) = \mathrm{P}(3) + \mathrm{P}(4) + \mathrm{P}(5) = \mathrm{C}_3^5 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} + \mathrm{C}_4^5 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} + \mathrm{C}_5^5 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} = 10 \cdot \frac{4}{243} + 5 \cdot \frac{2}{243} + 1 \cdot \frac{1}{243} = \frac{51}{243} = 0{,}209 = 21\%$

c) $\mathrm{P}(\text{5 respuestas correctas en 5 ítems}) = \mathrm{C}_5^5 \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} = 1 \cdot \frac{1}{243} = 0{,}004 = 0{,}4\%$

d) $\mathrm{P}(\text{0 respuestas correctas en 5 ítems}) = \mathrm{C}_0^5 \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = 1 \cdot \frac{32}{243} = 0{,}131 = 13{,}1\%$

Posteriormente, el estudiante requiere desarrollar el procedimiento de calcular la media y desviación estándar de la binomial mediante expresiones algebraicas. De esta manera: e) $\mu = \mathrm{n} \cdot \mathrm{p} = 5 \cdot \frac{1}{3} = 1{,}6$; y f) $\sigma = \sqrt{\mathrm{n} \cdot \mathrm{p} \cdot \mathrm{q}} = \sqrt{5 \cdot \frac{1}{3} \cdot \frac{2}{3}} = 1{,}05.$

Los diversos objetos matemáticos involucrados en las anteriores prácticas desarrolladas para resolver el I_9 e I_6 son presentados en la Tabla 14.

Tabla 14

Objetos matemáticos involucrados en la solución del I_9 e I_6

Tipo de objeto matemático	Objeto matemático	Ítem 9	6.a	6.b	6.c	6.d	6.e	6.f
Situaciones-problemas	Identificar situaciones que pueden modelarse a través de una distribución binomial	X	X	X	X	X	X	X
	Determinar los parámetros asociados a una distribución binomial		X	X	X	X	X	X
	Calcular la media y desviación estándar de la distribución binomial.						X	X
	Determinar probabilidades asociadas a una distribución binomial de forma manual		X	X	X	X		
Lenguaje	Verbal	X	X	X	X	X	X	X
	Numérico		X	X	X	X	X	X
	Simbólico	X	X	X	X	X	X	X
Conceptos	Experimento aleatorio	X	X	X	X	X	X	X
	Espacio muestral	X	X	X	X	X	X	X
	Suceso aleatorio	X	X	X	X	X	X	X
	Sucesos independientes	X	X	X	X	X		
	Variable aleatoria	X	X	X	X	X	X	X
	Enfoque de probabilidad clásico	X	X	X	X	X	X	X
	Distribución binomial	X	X	X	X	X	X	X
	Función de probabilidad de la binomial	X	X	X	X	X		
	Combinatoria	X	X	X	X	X		
	Media						X	
	Desviación estándar							X
Proposiciones	Caracterización de variables aleatorias discreta y continua	X	X	X	X	X	X	X
	Propiedades de la función de probabilidad		X	X	X	X		
	Caracterización de la distribución binomial	X	X	X	X	X		X
	La probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento	X	X	X	X	X	X	X
	Regla del producto		X	X	X	X		
	Regla de Laplace	X	X	X	X	X		
	La media es un representante de los datos a partir de los que ha sido calculada y no tiene por qué ser igual a uno de los valores de los datos						X	
Procedimientos	Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial	X	X	X	X	X		
	Cálculo de la probabilidad (de éxito) mediante la regla de Laplace	X	X	X	X	X		
	Cálculo de probabilidad compuesta con la regla del producto		X	X	X	X		
	Cálculo de combinaciones	X	X	X	X	X		
	Cálculo de media y desviación estándar de una binomial mediante sus expresiones algebraicas						X	X
Argumentos	Verbal-deductivo	X	X	X	X	X	X	X

Nota. Elaboración propia.

4.3.7 C-P_7 La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real

El C-P_7 está integrado por cuatro situaciones-problema (ver Figuras 17, 18, 19, 20 y 21): el $I_{10.b}$ e $I_{7.1}$ evalúan la S-$P_{7.1}$ identificar situaciones que pueden modelarse a través de una distribución normal (Bizet, et al. 2023c, p.187); el $I_{7.2}$ representa la S-$P_{7.2}$ determinar probabilidades asociadas a una distribución normal de forma manual (Bizet, et al. 2023c, p.187); el $I_{7.3}$ representa la S-$P_{7.3}$ describir la tendencia de los datos representados gráficamente empleando la aproximación de la distribución binomial por la normal (Bizet, et al. 2023c, p.187); y el $I_{7.4}$ valoran la S-$P_{7.4}$ evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal (Bizet, et al. 2023c, p.187).

Figura 17

S-$P_{7.1}$ que compone el C-P_7

10. Lee el siguiente titular, luego responda las preguntas propuestas

b) Considerando el gráfico expuesto ¿Podría ser verdadero el titular "siempre el 16% (de los estudiantes) obtiene más de 600 puntos en la PSU"? ¿Por qué?

Nota. Elaboración propia y posee información de Ramírez (2018).

Para responder el $I_{10.b}$ el estudiante requiere identificar que los puntajes obtenidos por los individuos en la prueba PSU son normalizados, donde se utiliza una escala de puntaje estándar con media de 500 puntos y desviación estándar de 110 puntos, lo que implica que: i) entre los 390 y 610 puntos hay un determinado porcentaje de individuos, que es 68,26%; ii) entre 280 y 720 hay un 95,44%; y iii) entre 170 y 830 puntos se sitúa el 99,72% de la población. Luego $\frac{99,72\%-68,26\%}{2} = 15,73\%$, entonces es verdadera la afirmación propuesta ya que alrededor de un 16% de las personas que rinden la prueba obtienen un puntaje mayor o igual a 610 puntos.

Figura 18

S-P

S-$P_{7.1}$ que compone el C-P_7

7. 1 ¿En cuál de los siguientes casos la variable sigue una distribución normal?
I. Lanzamiento de un dado.
II. Estatura de una población.
III. Lanzamiento de una moneda.
A) Solo I.
B) Solo II.
C) Solo III.
D) I y II.

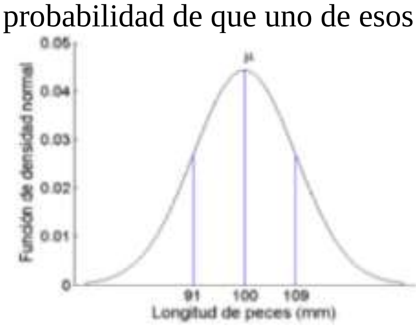
Nota. Extraído de Morales et al. (2008, p.457).

Sobre la resolución al $I_{7.1}$ el estudiante debe recordar características de la distribución normal, es decir, las cualidades del contexto donde esta distribución es aplicable*:* (i) situación modelada por una variable aleatoria continua; y (ii) sus valores se agrupan en torno a un valor central (media) y sus valores extremos no son frecuentes. Luego si este observa la variable aleatoria continua estatura de compañeros de asignatura elegidos al azar, notará que la mayoría tiene estatura media del grupo, y si ve a alguien que se aparta mucho de la media llama su atención, así puede reconocer que II es verdadera. También este requiere reconocer que para los experimentos aleatorios de I e II solo se pueden definir variables aleatorias discretas, entonces la respuesta corresponde a la alternativa B.

Figura 19

S-$P_{7.2}$ y S-$P_{7.3}$ que compone el C-P_7

7.2 La longitud de cierto tipo de peces sigue una distribución normal con media $\mu = 100$ y varianza $\sigma^2 = 81$. ¿Cuál es la probabilidad de que uno de esos peces mida entre 82 y 91mm?



Nota. Extraído de González y Ojeda (2017, p.209).

Para resolver el $I_{7.2}$ el estudiante primero debe identificar que la situación propuesta está modelada por X~N(100,9) donde la variable aleatoria continua X es la longitud de un pez, luego se espera que desarrolle el procedimiento de estandarización, transformar la variable aleatoria X~N(100,9) en Z~N(0,1) así $Z_1 = \frac{x_1-\sigma}{\mu} = \frac{82-100}{9} = -2$ y $Z_2 = \frac{x_2-\sigma}{\mu} = \frac{91-100}{9} = -1$, entonces obtienen que $P(82 \leq X \leq 91) = P(Z \leq -1) - P(Z \leq -2)$. Posteriormente este requiere aplicar propiedades para el cálculo de probabilidades con distribución normal estándar, de esta manera consigue que $P(Z \leq -1) - P(Z \leq -2) = P(Z \geq 1) - P(Z \geq 2) = 1 - P(Z < 1) - [1 - P(Z < 2)] = P(Z < 2) - P(Z < 1)$, finalmente necesita desarrollar los procedimientos de calcular probabilidades utilizando la función de densidad de la normal estándar y leer su tabla de probabilidades, es decir, $P(Z < 2) - P(Z < 1) = 0{,}9772 - 0{,}8413 = 0{,}1359$. Por tanto $P(82 \leq X \leq 91) = 0{,}1359$ o 13,59%

Otra estrategia que el estudiante puede realizar es aplicar una propiedad de la función de densidad de la normal, específicamente debe utilizar la vinculada a los intervalos centrales ($\pm 3\sigma$), o sea, $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0{,}683$ y $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0{,}954$. Así $P(82 \leq X \leq 91) = \frac{P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) - P(\mu - \sigma \leq X \leq \mu + \sigma)}{2} \approx \frac{0{,}954 - 0{,}683}{2} \approx 0{,}1355$

Figura 20

S-P$_{7.4}$ que compone el C-P$_{7}$

7.3 Observa las siguientes distribuciones binomiales (Figura 1) ¿Qué sucede con el gráfico de barras cuando el n crece? Anota tu explicación.

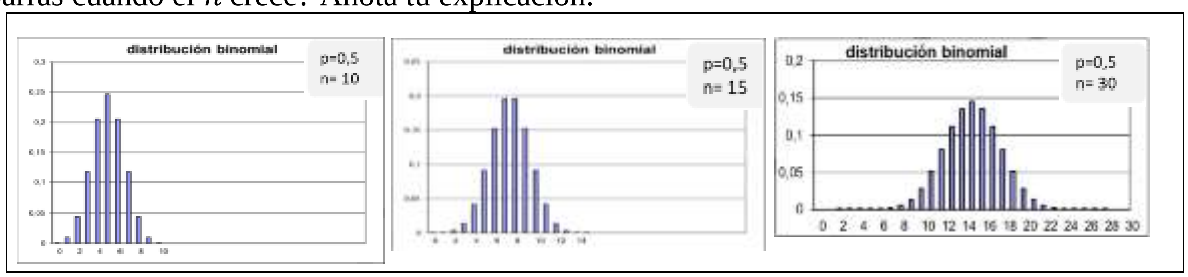


Figura 1. Secuencia de distribuciones binomiales

Nota. Adaptación de MINEDUC (2019c, p.56).

Sobre la resolución del I$_{7.3,}$ inicialmente el estudiante debe leer e interpretar los gráficos de barras propuestos, para establecer que a medida que crece el n número veces que se repite el experimento aleatorio, el gráfico de barras de la distribución binomial (con p=0,5 probabilidad de éxito en un intento) se acerca a la curva en forma de campana de la distribución normal. Luego este necesita utilizar la condición para aproximar una distribución binomial a una normal, así determinar que se cumple: (i) $n = 30 \geq 30$; (ii) $n \cdot p = 15 > 5$; y (iii) $n \cdot (1 - p) = 15 > 5$, entonces $B(30; 0{,}5) \sim N(15; 2{,}7)$.

Figura 21

S-$P_{7.5}$ que compone el C-P_7

7.4 V/F Justifica tu respuesta "En una distribución normal, el 50% de las medidas caen por encima de la media".

Nota. Extraídos de Tauber (2001, p.182)

La resolución del $I_{7.4}$ requiere que el estudiante emplee alguna propiedad de la función de densidad de la normal: i) esta función es simetría con respecto a su media, la cual significa que $P(x \geq \mu) = P\,(x \leq \mu) = 0{,}5 = 50\%$; o ii) la media es, además, la mediana de la distribución normal. En consecuencia, la afirmación presentada es verdadera.

La Tabla 15 muestra los objetos matemáticos que intervienen en las prácticas anteriormente descritas desarrolladas para resolver el $I_{10.b}$ e $I_{7.}$

Tabla 15

Objetos matemáticos involucrados en la solución del $I_{10.b}$ e I_7

Tipo de objeto matemático	Objeto matemático	Ítem 10.b	7.1	7.2	7.3	7.4
Situaciones-problemas	Identificar situaciones que pueden modelarse a través de una distribución normal	X	X			
	Determinar probabilidades asociadas a una distribución normal de forma manual			X		
	Describir la tendencia de los datos representados gráficamente empleando la aproximación de la distribución binomial por la normal				X	
	Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal					X
Lenguaje	Verbal	X	X	X	X	X
	Numérico			X		
	Simbólico			X	X	X
	Gráfico	X		X	X	
	Tabular			X		
Conceptos	Experimento aleatorio		X		X	
	Intervalo	X		X		
	Variable aleatoria	X	X	X	X	
	Distribución binomial				X	
	Enfoque de probabilidad clásico			X		
	Media	X	X	X	X	X
	Mediana					X
	Varianza			X		
	Desviación estándar	X		X	X	
	Distribución normal	X	X	X	X	X
	Función de densidad de la normal			X		X
	Simetría	X		X		X
	Distribución normal estándar			X		
Proposiciones	Caracterización de variables aleatorias discreta y continua	X	X	X	X	
	Caracterización de la distribución normal	X	X	X	X	
	Propiedades de la función de densidad de la normal					X

	Propiedad de los intervalos centrales (±3σ)	X		X		
	Propiedades para el cálculo de probabilidades con distribución normal estándar			X		
	Condición para aproximar una distribución binomial a una normal				X	
Procedimientos	Lectura e interpretación de un gráfico de barras				X	
	Lectura e interpretación de una curva normal	X				
	Verificación de propiedades o característica de la distribución normal		X		X	X
	Cálculo de probabilidad empleando la propiedad ±3σ	X		X		
	Tipificación			X		
	Cálculo de probabilidades utilizando la función de densidad de la normal estándar			X		
	Lectura de tabla de la función de densidad de la normal estándar			X		
	Verificación de condiciones para la aproximación a una distribución normal				X	
Argumentos	Verbal-deductivo		X	X		X
	Mediante representación gráfica	X			X	

Nota. Elaboración propia.

4.3.8 C-P$_8$ Aproximar distribuciones de variables aleatorias discretas con un número grande de valores

El C-P$_8$ está constituido por dos situaciones-problema (ver Figura 22 y 23): el I$_{8.1}$ evalúa la S-P$_{8.1}$ calcular los parámetros asociados a una distribución normal como media μ y desviación estándar σ (Bizet, et al. 2023c, p.187); y el I$_{8.2}$ valora la S-P$_{8.2}$ calcular probabilidades en el contexto de aproximación de la distribución binomial a la normal (Bizet, et al. 2023c, p.187).

Figura 22

S-P$_{8.1}$ que compone el C-P$_8$

8.1 El 4% de los clavos de 2 pulgadas producidos por una empresa salen defectuosos. Si se produce un lote de 1200 clavos, determina el ajuste a la normal.

Nota. Extraído de Norambuena et al. (2019, p.43).

Para responder al I$_{8.1}$, el estudiante necesita definir la variable aleatoria Y: cantidad de clavos defectuosos en una muestra de tamaño 1200, e identificar que $Y \sim B(1200; 0{,}04)$. Posteriormente este requiere comprobar que se cumplen la condición para aproximar una distribución binomial a una normal: i) $n = 1200 \geq 30$; ii) $n \cdot p = 1200 \cdot 0{,}04 = 48 > 5$; y iii) $n(1-p) = 1200 \cdot 0{,}96 = 1152 > 5$, entonces puede determinar que $Y \sim (1200; 0{,}04)$ se puede aproximar a $X \sim N(48; 6{,}79)$, donde $\mu = 1200 \cdot 0{,}04 = 48$ y $\sigma = \sqrt{1200 \cdot 0{,}04 \cdot 0{,}96} = \sqrt{46{,}08} = 6{,}79$.

Figura 23

S-P$_{8.2}$ que compone el C-P$_8$

8.2 Sean $X_i = 0$ si un producto es defectuoso y $X_i = 1$ si el producto está correcto. En un lote de 30 de estos productos queremos calcular la probabilidad de que 25 al menos sean correctos, aproximando por la normal. La corrección de continuidad para el cálculo implica que tenemos que calcular cuál de las siguientes expresiones: A) $P(\sum_{i=1}^{30} X_i \geq 25{,}5)$ B) $P(\sum_{i=1}^{30} X_i \geq 24{,}5)$ C) $P(\sum_{i=1}^{30} X_i \geq 25)$ D) $P(\sum_{i=1}^{30} X_i \leq 25{,}5)$

Nota. Extraído de Alvarado y Batanero (2007, p.4).

En la solución al I$_{8.2}$, el estudiante debe reconocer que $X \sim B(30; 0{,}5)$ y se cumplen la condición para aproximar aquella distribución binomial a una normal: i) $n = 30 \geq 30$; ii) $n \cdot p = 30 \cdot 0{,}5 = 45 > 5$; y iii) $n(1-p) = 30 \cdot 0{,}5 = 45 > 5$, entonces $B(30; 0{,}5) \sim N(45; 2{,}7)$. Luego para calcular la probabilidad de la variable aleatoria discreta X mediante la aproximación de la normal, este necesita aplicar la corrección por continuidad, $P(\sum_{i=1}^{30} X_i \geq 25 - 0{,}5) = P(\sum_{i=1}^{30} X_i \geq 24{,}5)$, así la respuesta es la alternativa B.

La Tabla 16 presenta los objetos matemáticos movilizados en las anteriores prácticas llevadas a cabo para resolver el I$_8$.

Tabla 16

Objetos matemáticos involucrados en la solución del I$_8$

Tipo de objeto matemático	Objeto matemático	Ítem 8.1	Ítem 8.2
Situaciones-problemas	Calcular los parámetros asociados a una distribución normal como media μ y desviación estándar σ	X	
	Calcular probabilidades en el contexto de aproximación de la distribución binomial a la normal		X
Lenguaje	Verbal	X	X
	Numérico	X	
	Simbólico	X	X
Conceptos	Variable aleatoria	X	X
	Distribución binomial	X	X
	Media	X	X
	Desviación estándar	X	X
	Distribución normal	X	X
Proposiciones	Caracterización de variables aleatorias discreta y continua	X	X
	Caracterización de la distribución binomial	X	X
	Condición para aproximar una distribución binomial a una normal	X	X
Procedimientos	Cálculo de media y desviación estándar mediante expresiones algebraicas	X	X
	Verificación de condiciones para la aproximación a una distribución normal	X	X
	Corrección por continuidad		X
Argumentos	Verbal-deductivo	X	X

Nota. Elaboración propia.

4.4 Estudio 5[6] Cuestionario para valorar la comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar

Bizet, V., Molina-Portillo, y Contreras, J.M. (en evaluación). Cuestionario para valorar la comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar.

Resumen

Esta investigación presente una propuesta de cuestionario que valora la comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar chilena, fundamentado en el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos. El objetivo de la investigación es analizar la validez de constructo y fiabilidad de dicho instrumento. Aquella se aborda desde un enfoque cuantitativo y es de tipo descriptiva-correlacional, donde la versión inicial del instrumento se aplicada en una muestra dirigida de 80 estudiantes. La validez de constructo del instrumento es indagada mediante un análisis factorial exploratorio (AFE) y confirmatorio (AFC), y su fiabilidad empleando el coeficiente de alfa de Cronbach. Los resultados muestran que, en el AFE, la estructura empírica del cuestionario está compuesta por seis factores que explican el 58% de la varianza total, cada uno relacionado con un campo de problema (constituido por tres a cuatro ítems) sobre los temas en cuestión: variable aleatoria, función de probabilidad y función de distribución de una variable aleatoria discreta, valores de posición central o de dispersión asociados a una variable aleatoria, distribuciones binomial y normal. En el AFC, la mayoría de los coeficientes factor-ítem poseen valores adecuados y la estructura del instrumento tiene medidas de ajuste aceptables, además el alfa de Cronbach de 0,824 demuestra que el cuestionario es fiable y existe consistencia entre los ítems y constructo. En conclusión, la versión final del cuestionario (19 ítems) posee evidencias que es un instrumento válido y fiable para indagar la compresión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar chilena.

---

[6] Para evitar confusión y repetición en la numeración, se ha modificado la relativa a los epígrafes pertenecientes a los estudios que forman parte del compendio, haciendo referencia al capítulo en que se ubica cada estudio. También se ha cambiado la numeración de Tablas y Figuras con respecto a la versión publicada del artículo, haciendo referencia al estudio en que se enmarcan.

Palabras clave: variable aleatoria, distribuciones de probabilidad, evaluación, educación escolar.

Abstract

This paper proposes a questionnaire to assess the understanding of random variables and probability distributions in Chilean school graduates, according to the Ontosemiotic Approach to Mathematical Knowledge and Mathematical Instruction. This study analyses the construct validity and reliability of the instrument. The research is quantitative and descriptive-correlational, where the initial version of the instrument was applied to a sample of 80 students. The validity of the instrument's construct is tested through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) and the reliability using Cronbach's alpha coefficient. The results show that, in the AFE, the empirical structure of the questionnaire is composed of six factors that explain 58% of the total variance. Each factor is related to a problem domain (consisting of three to four items) on the topics in question: random variable, probability function and distribution function of a discrete random variable, measures of central tendency or measures of dispersion associated to a random variable, binomial and normal distributions. In the CFA, most factor-item coefficients have good values and the instrument's structure has acceptable fit measures. Furthermore, Cronbach's alpha of 0,824 shows that the questionnaire is reliable and consistent between items and construct. To conclude, there is evidence that this questionnaire (19 items) is a valid and reliable instrument to investigate the understanding of random variables and probability distributions in Chilean school leavers.

Keywords: random variable, probability distributions, assessment, school education.

4.4.1 Introducción

En las últimas dos décadas, el currículo escolar de diversos países ha incluido progresivamente el estudio de la probabilidad a lo largo de sus 12 grados (6-18 años) de extensión. Así su implementación ha generado un escenario desafiante para la didáctica de la estadística debido a la complejidad conceptual más que operacional de diversos conceptos probabilísticos (Del Pino y Estrella, 2012). Específicamente la variable aleatoria y su familia de distribuciones de probabilidad son conceptos abordados en diversos lineamientos curriculares internacionales dirigidos a los grados 9 a 12 (14-18 años) y su relevancia de estudiarlos en la escuela se debe a que se manifiestan en situaciones cotidianas (Pfannkuch, 2018).

Un ejemplo es lo realizado por Estados Unidos, que entre los grados 9 a 12 incluye el estudio de variable aleatoria, su media y su función de probabilidad, y la distribución binomial, en los Principios y Estándares para la Matemática Escolar (NCTM, 2000), además de la distribución normal en los Estándares Estatales Comunes de Matemática (NGACBP y CCSSO, 2010). Chile no queda ajeno a este escenario, en las Bases Curriculares (MINEDUC, 2015; 2019a) en grado 10 (15-16 años) introduce la variable aleatoria y su función nombrada, para grado 11 (16-17 años) su media, varianza y desviación estándar, y en grado 12 (17-18 años) las distribuciones binomial y normal.

Sin embargo, son limitadas las investigaciones centradas en la comprensión de variable aleatoria y distribuciones de probabilidad a nivel escolar (Bizet et al., 2022), entonces se desconoce cuáles son los conocimientos, principales conflictos de aprendizaje y dificultades de los estudiantes al terminar su etapa educativa obligatoria. Por tanto, el objetivo de esta investigación es analizar la validez de constructo y fiabilidad de un instrumento que valora la comprensión sobre variable aleatoria y sus distribuciones de probabilidad en estudiantes egresados de educación escolar chilena, fundamentado en el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos (EOS) (Godino et al., 2020).

4.4.2 Antecedentes

En relación con los estudios previos más importantes sobre el aprendizaje de variable aleatoria a nivel escolar, se han encontrado investigaciones en torno a comprender la variable aleatoria discreta y su función de probabilidad (Bizet y Ramos, 2022; Doukhan y Gueudet, 2019; García et al., 2014; García y Sánchez, 2013), además de comprender algunos valores de posición central o de dispersión vinculados a ella, como la media (Guerrero et al., 2017) y desviación estándar respectivamente (Armah et al., 2016). En este contexto, en el razonamiento de estudiantes se reconoce la presencia del sesgo cognitivo de resultados aislados (Flores et al., 2014). También son evidenciadas dificultades relativas a: identificar el dominio (Bizet y Ramos, 2022) y recorrido (García et al., 2014) de una variable aleatoria discreta, reconocerla como una función (Salazar, 2014), y calcular su media (Guerrero et al., 2017). Las investigaciones existentes relativas a comprender la variable aleatoria continua, solo se han desarrollado a nivel universitario.

Entre los estudios previos que involucran el aprendizaje de distribuciones de probabilidad en el contexto escolar, se han identificado investigaciones sobre comprender la distribución binomial (Bill et al., 2009; Van Dooren, et al., 2003) y su media (Begué et al., 2020). En aquel ámbito son identificados tanto los sesgos cognitivos de equiprobabilidad (Sánchez et al., 2018) y linealidad (Van Dooren et al., 2003), como dificultades en torno a aplicar su función de probabilidad (Sánchez y Landín, 2014) y su función de distribución (Sánchez y Carrasco, 2018) para calcular probabilidades. Además, se han reconocido estudios relacionados con comprender la distribución normal (González y Ojeda, 2017; Shin, 2012; Valdez y Salinas, 2019), contexto en el cual es reconocido el sesgo de insensibilidad al tamaño de la muestra (Alvarado et al., 2018), y dificultades para entender etapas del procedimiento de estandarización y relacionar el área bajo la curva normal con la probabilidad (Valdez y Salinas, 2019).

Esta revisión ha permitido reconocer que las investigaciones existentes emplean cuestionarios para evaluar la comprensión de variable aleatoria, distribución binomial o distribución normal sin indicar los procesos de validación de los instrumentos (validez de contenido y/o validez de constructo). Además, se ha identificado la carencia de herramientas que valoren articulada y simultáneamente la comprensión de dichos temas. Aquel escenario se agrava frente al inminente estudio de estos contenidos durante la transición de término de la educación escolar y comienzo de la educación universitaria, que se espera sea congruente con las normativas curriculares actuales. Por tanto, la elaboración de un instrumento que recopile datos válidos y fiables sobre la comprensión de variable aleatoria y distribuciones de probabilidad al finalizar la escuela es una demanda urgente para obtener información orientada a fortalecer la educación estadística.

4.4.3 Fundamentación

4.4.3.1 Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos

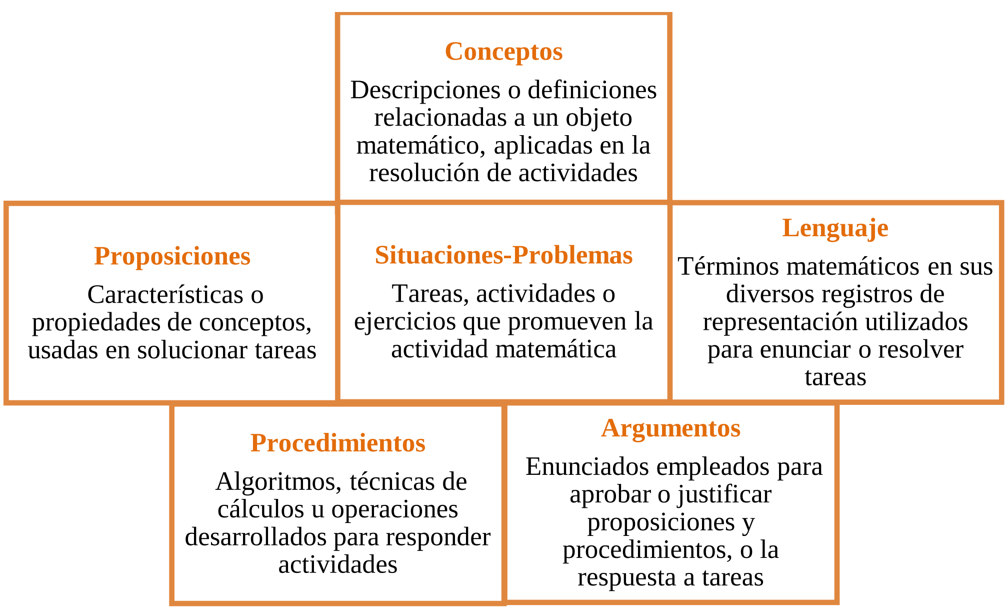
La presente investigación se fundamenta en el EOS (Godino et al., 2020), teoría que se origina de las ideas de situaciones-problemas (tareas, actividades o problemas) y prácticas matemáticas realizadas para su resolución. Está ultima se refiere a específicas actuaciones y expresiones verbales o gráficas realizadas por una persona o compartidas en una institución para resolver una tarea o comunicar su solución (Godino, 2017). Las prácticas matemáticas atribuidas a una persona son denominadas significado personal de un objeto matemático y aquellas llevadas a cabo dentro de una institución son llamadas significado institucional.

Dentro del significado institucional de un objeto matemático es posible distinguir cuatro tipos (Godino et al., 2007): i) el referencial, usado como pauta para elaborar el significado a trabajar en el aula, presente en libros de texto científicos y lineamientos curriculares, entre otros; ii) el pretendido, integrado en la planificación de la enseñanza; iii) el implementado, aplicado de manera efectiva por el profesor en el aula; y iv) el evaluado, empleado por el profesor o investigador para evaluar los aprendizajes en torno al objeto. Mientras que en el significado personal se identifican tres tipos (Godino et al., 2007): i) el global, relativo a las prácticas que potencialmente el estudiante podría lograr sobre el objeto matemático específico; ii) el declarado, corresponde al plasmado por el estudiante en evaluaciones o tareas, el cual puede ser adecuado o erróneo; y iii) el logrado, referido a las prácticas realizadas por el estudiante sobre un objeto concreto y coherentes con las propuestas por la institución.

Para el EOS los objetos corresponden a todo aquello a lo que se hace referencia en la práctica matemática, por lo que según la función que cumplan es posible clasificarlos en la tipología de objetos matemáticos primarios (Godino, 2002) expuesta en la Figura 5.1. Así, un conjunto de situaciones-problemas vinculadas recíprocamente, que comparten sus representaciones, procesos o soluciones similares constituyen un campo de problema (Godino, 1999, p. 199).

Figura 5.1

Tipos de objetos matemáticos primarios



Nota. Elaboración propia.

En términos del EOS la comprensión (Godino, 2002) no es observable directamente en el estudiante, sino que se deduce de la observación a las respuestas de los participantes a situaciones-problemas, por lo que la valoración de la comprensión se entiende como la correspondencia entre el significado personal y significado institucional de un objeto matemático. De esta manera, este estudio emplea tanto el significado institucional evaluado en el instrumento diseñado como el significado personal declarado por los estudiantes.

4.4.3.2 Campos de problemas de variable aleatoria y distribuciones de probabilidad

Para identificar los principales campos de problemas que dotan de sentido al proceso de enseñanza-aprendizaje de la variable aleatoria y distribuciones de probabilidad, fue desarrollado el procedimiento realizado por Alvarado (2007) y Gea (2014) en investigaciones sobre educación matemática. Este consiste en la indagación de estudios previos llevados a cabo desde el ámbito epistemológico, que abarcan el origen del concepto, su desarrollo histórico y estudio desde la disciplina matemática.

Respecto a la variable aleatoria, las investigaciones en torno a su aspecto epistemológico desarrolladas por Dinges (2005), Ruiz (2013) y Amrani y Zaki (2015) permitieron establecer los siguientes campos de problemas (Bizet et al., 2023a, p.175): C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico; C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria; C-P_3 Reconocer la distribución de probabilidad de una variable aleatoria continua como una herramienta que permite ver la variación aleatoria; C-P_4 Establecer la función de distribución de una variable aleatoria discreta; y C-P_5 Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria.

Sobre la distribución binomial, su estudio desde la matemática realizado por Vilca (2015) y la investigación relativa a su origen y evolución histórica llevada a cabo por García-García et al. (2022) permitieron determinar el C-P_6 La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real (Bizet et al., 2023a, p.175). Respecto a la distribución normal, su indagación a partir de la disciplina matemática desarrollada por Tauber (2001) y el trabajo en torno a su génesis y desarrollo histórico realizado por Stahl (2006) posibilitaron identificar el C-P_7 La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real, y el C-P_8 Aproximar distribuciones de variables aleatorias discretas con un número grande de valores (Bizet et al., 2023a, p.175).

4.4.4 Metodología

Este estudio fue abordado desde un enfoque cuantitativo, es de tipo descriptivo-correlacional y utilizó un diseño transversal (Hernández et al., 2014).

4.4.4.1 Instrumento

El instrumento evalúa la comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar chilena (Bizet et al., 2023a) y su construcción abarcó tres fases. En la primera fase, fue creada la guía de situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad según el currículo escolar chileno (GSP-VADP). Para ello, inicialmente se establecieron los principales campos de problemas de variable aleatoria y distribuciones binomial y normal, con base en el análisis epistemológico desarrollado previamente (Bizet et al., 2022). Luego fue desarrollado un análisis de contenido del currículo escolar chileno para los grados 10 a 12 (MINEDUC, 2015; 2019a) y lineamientos curriculares internacionales sobre el tratamiento otorgado a los temas en cuestión (NTCM, 2000; NGACBP y CCSSO, 2010). A partir de este se infirieron las situaciones-problemas que integran la GSP-VADP.

En la segunda fase, desde la literatura se eligió un conjunto inicial de ítems que componen el instrumento de evaluación. Puntualmente para cada situación-problema de la GSP-VADP se propusieron tres ítems, que fueron seleccionados de investigaciones previas, tareas propuestas en libros de texto previamente analizados (Bizet et al., 2023b) o elaborados por los autores. En la tercera fase, fue elegido el conjunto final de ítems que componen el instrumento, mediante la validez de contenido por juicio de expertos. Participaron seis doctores en educación matemática con experiencia en investigación en didáctica de la estadística, quienes evaluaron de forma independiente tres ítems para cada tarea. Así se obtuvo la versión inicial del instrumento, seleccionándose en cada situación-problema el ítem con media más alta y menor desviación estándar.

El cuestionario se organiza en ocho campos de problemas (C-P), cada uno compuesto por dos a seis situaciones-problemas (S-P), donde los ítems (I) que valoran a cada una de ellas se exponen en la Tabla 5.1.

Tabla 5.1

Estructura teórica del instrumento según la GSP-VADP

C-P	Situación-Problema	Ítem
C-P$_1$	S-P$_{1.1}$ Diferenciar entre variables aleatorias y variables con dependencia funcional	I$_{1.1}$
	S-P$_{1.2}$ Definir variables aleatorias finitas en experimentos aleatorios	
	S-P$_{1.3}$ Identificar dominio de una variable aleatoria finita	I$_{1.2}$
	S-P$_{1.4}$ Identificar recorrido o valores de una variable aleatoria finita	I$_{1.3}$
C-P$_2$	S-P$_{2.1}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial	I$_{2.1a}$
	S-P$_{2.2}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico	I$_{2.2a}$
	S-P$_{2.3}$ Definir la función de probabilidad de una variable aleatoria utilizando la terminología P (X = x$_i$)	I$_{2.2b}$
	S-P$_{2.4}$ Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta	I$_{2.1b}$
	S-P$_{2.5}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta	I$_{2.2c}$
C-P$_3$	S-P$_{3.1}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua	I$_{3.1}$
	S-P$_{3.2}$ Calcular algunas probabilidades asociadas a una variable aleatoria continua	I$_{3.2b}$
	S-P$_{3.3}$ Determinar si la función dada es una función de probabilidad de una variable aleatoria continua	I$_{3.2a}$
C-P$_4$	S-P$_{4.1}$ Determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta	I$_{4.a}$
	S-P$_{4.2}$ Definir la función de distribución de una variable aleatoria discreta	
	S-P$_{4.3}$ Representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta	I$_{4.b}$
C-P$_5$	S-P$_{5.1}$ Calcular la media o valor esperado de una variable aleatoria discreta	I$_{5.1}$
	S-P $_{5.2}$ Interpretar la media o valor esperado de una variable aleatoria discreta	I$_{5.2c}$
	S-P $_{5.3}$ Calcular la varianza de una variable aleatoria discreta	I$_{5.2b}$
	S-P $_{5.4}$ Calcular la desviación estándar de una variable aleatoria discreta	I$_{5.2a}$
	S-P$_{5.5}$ Interpretar la desviación estándar de una variable aleatoria discreta	I$_{5.2c}$
	S-P$_{5.6}$ Identificar la media y desviación estándar de una variable aleatoria continua	I$_{10.a}$
C-P$_6$	S-P$_{6.1}$ Identificar situaciones que pueden modelizarse a través de una distribución binomial	I$_9$ e I$_6$
	S-P$_{6.2}$ Determinar los parámetros asociados a una distribución binomial como n (número de ensayos), p (probabilidad de éxito) y q (probabilidad de fracaso)	I$_6$
	S-P$_{6.3}$ Calcular la media y desviación estándar de una distribución binomial	I$_{6.e}$ e I$_{6.f}$
	S-P$_{6.4}$ Determinar probabilidades asociadas a una distribución binomial de forma manual	I$_{6.a}$, I$_{6.b}$, I$_{6.c}$ e I$_{6.d}$
C-P$_7$	S-P$_{7.1}$ Identificar situaciones que pueden modelizarse a través de una distribución normal	I$_{7.1}$ e I$_{10.b}$
	S-P$_{7.2}$ Determinar probabilidades asociadas a una distribución normal de forma manual	I$_{7.2}$
	S-P$_{7.3}$ Describir la tendencia de datos representados gráficamente empleando la aproximación de la binomial por la normal	I$_{7.3}$
	S-P$_{7.4}$ Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal	I$_{7.4}$
C-P$_8$	S-P$_{8.1}$ Calcular los parámetros asociados a una distribución normal como μ (media) y σ (desviación estándar)	I$_{8.1}$

S-$P_{8.2}$	Calcular probabilidades en el contexto de aproximación de la distribución binomial a la normal	$I_{8.2}$

Nota. Extraída de Bizet et al. (2023a).

Fue desarrollado un análisis a priori de los ítems para identificar los objetos matemáticos primarios involucrados en las situaciones-problemas que representan. Así aquel conjunto de prácticas permitió identificar el significado institucional evaluado en la versión inicial del cuestionario.

4.4.4.2 Participantes

La población de interés fue estudiantes egresados de educación escolar chilena, por lo que se utilizó una muestra dirigida (no probabilística) de 80 estudiantes universitarios chilenos quienes no habían cursado en la universidad alguna asignatura sobre estadística y probabilidad. Puntualmente colaboraron 31 estudiantes de primer año de la carrera de Pedagogía en Educación Especial y 49 estudiantes de la carrera de Pedagogía en Matemática (8 de primer año, 28 de segundo y 13 de tercero). Los participantes respondieron el cuestionario voluntaria e individualmente en una sesión de 90 minutos (entre agosto y septiembre de 2021), bajo la modalidad en línea por motivo de la Covid-19. Es importante señalar que, en investigaciones sobre educación para analizar la fiabilidad y validez de constructo de instrumentos se han empleado muestras de tamaño pequeño, por ejemplo, Martínez (2020) utilizó 98 inspectores de educación.

4.4.4.3 Procedimientos

Para el análisis de las respuestas proporcionadas por los participantes se utilizó la técnica de análisis de contenido (Krippendorff, 1990). Las respuestas a los ítems fueron clasificadas en tres categorías y se asignó puntuación a cada una: respuesta correcta dos puntos, parcialmente correcta un punto (algunos objetos matemáticos vinculado a una situación-problema son puesto en práctica para su resolución); errónea cero puntos, las respuestas en blanco fueron consideradas en esta categoría.

Luego se indagó la validez de constructo, aquella fue posible establecer a través del análisis factorial y verificación de la teoría subyacente (Hernández et al., 2014), por lo que se realizaron los siguientes procedimientos:

1. Fueron realizadas las pruebas de idoneidad de los datos para el análisis factorial, mediante la medida de adecuación muestral de Kaiser, Meyer, Olkin que es suficiente si $0{,}70 \leq$

$\mathrm{KMO} \leq 0{,}79$ (Lloret-Segura et al., 2014), y la prueba de esfericidad de Bartlett, en la cual se requiere que los resultados conseguidos sean significativos (p-valor<0,05) para rechazar la hipótesis nula y afirmar que las variables están interrelacionadas (Everitt y Wykes, 2001).

2. Se desarrolló una exploración de la estructura del cuestionario a través de un análisis factorial exploratorio (AFE) y llevó a cabo una aproximación a confirmar la estructura del instrumento por medio de un análisis factorial confirmatorio (AFC).

3. Fue evaluado el ajuste de la estructura del cuestionario presentada en el AFC, a través de algunas medidas de bondad de ajuste que deben cumplir las siguientes condiciones (Kline, 2005): el error de aproximación cuadrático medio, $\mathrm{RMSEA} \leq 0{,}08$; el índice de error cuadrático medio, $\mathrm{SRMR} \leq 0{,}08$; el índice no normalizado de ajuste, $\mathrm{TLI} \geq 0{,}9$; y el índice de Ajuste Comparativo, $\mathrm{CFI} \geq 0{,}9$.

Finalmente se estudió la fiabilidad del instrumento a través del método medidas de consistencia interna y empleando el coeficiente alfa de Cronbach (Hernández et al., 2014), el cual si toma un valor entre 0,8 y 0,9 se considera una buena consistencia (George y Mallery, 2003). Todos los procedimientos señalados anteriormente se realizaron mediante los softwares SPSS-Statistics (v25) y Stata (v15).

4.4.5 Resultados

4.4.5.1 Análisis factorial exploratorio

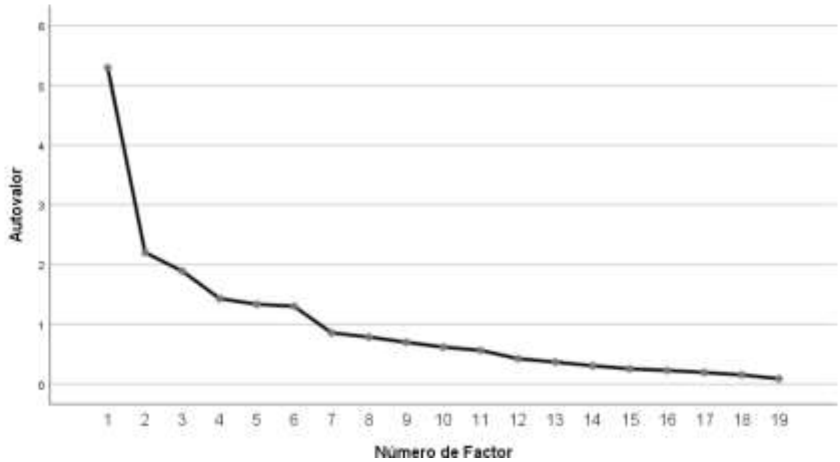
Fue desarrollada la prueba de Bartlett ($x^2 = 690{,}668$; $\mathrm{gl} = 171$; $\mathrm{p-valor} = 0{,}00$), la cual permitió afirmar que se rechaza la hipótesis nula y acepta la hipótesis alterna que las variables están interrelacionadas (la matriz de correlaciones no es una matriz identidad), además se calculó la medida de adecuación muestral KMO, obteniéndose un valor suficiente de 0,73. Por tanto, los resultados logrados garantizaron la pertinencia de realizar un AFE.

En el análisis factorial exploratorio para la estimación de factores se utilizó el método de ejes principales, debido a que no se cumple la normalidad de las puntuaciones (Lloret-Segura et al., 2014), y fue realizada una rotación ortogonal de la matriz de componentes principales. Así se han retenido seis factores (Ver Figura 5.2), donde cada valor propio es al menos 1 (Rodríguez-Santero y Gil-Flores, 2019), y aquellos explican el 58% de la varianza total (ver Tabla 5.2).

Previamente, de los 32 ítems se eliminaron 13($I_{1.2}$, $I_{3.2a}$, $I_{3.2b}$, $I_{5.1}$, $I_{5.2a}$, $I_{6.a}$, I_{9}, $I_{6.b}$, $I_{6.f}$, $I_{7.1}$, $I_{10.b}$, $I_{8.1}$ e $I_{8.2}$), porque en algún factor no presentaron una carga factorial mayor a 0,4 como recomiendan Williams et al. (2010).

Figura 5.2

Gráfico de sedimentación



Nota. Elaboración propia.

Tabla 5.2

Varianza total explicada por cada uno de los factores retenidos

Factor	Sumas de cargas al cuadrado de la rotación		
	Total	% de varianza	% acumulado
1	2,589	13,625	13,625
2	2,350	12,366	25,991
3	2,099	11,049	37,040
4	1,469	7,729	44,769
5	1,287	6,774	51,543
6	1,281	6,741	58,284

Nota. Elaboración propia.

La Tabla 5.3 expone los ítems que componen cada factor, además está incluye una columna relativa al campo de problema que pertenece cada ítem. Así la estructura empírica del instrumento está compuesta por seis factores. El primer factor se podría asociar al C-P_4 Establecer la función de distribución de una variable aleatoria discreta. El segundo factor relacionado con el C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria. El tercer factor agrupa los ítems del C-P_6 La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real. El cuarto factor se vincula al C-P_5 Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria. El quinto factor asociado al C-P_7 La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real. Por último, el sexto factor posible de vincularse al C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico

Tabla 5.3

Matriz de componentes rotados

C-P	Ítem	Factor					
		1	2	3	4	5	6
C-P_4	$I_{4.a}$	0,934					
C-P_2	$I_{2.2c}$	0,822					
C-P_4	$I_{4.b}$	0,682					
C-P_2	$I_{2.1b}$		0,837				
	$I_{2.1a}$		0,803				
	$I_{2.2a}$		0,529				
	$I_{2.2b}$		0,501				
C-P_6	$I_{6.d}$			0,896			
	$I_{6.c}$			0,823			
	$I_{6.e}$			0,524			
C-P_5	$I_{5.2c}$				0,757		
	$I_{10.a}$				0,568		
	$I_{5.2b}$				0,401		
C-P_7	$I_{7.3}$					0,521	
	$I_{7.2}$					0,512	
	$I_{7.4}$					0,468	
C-P_1	$I_{1.1}$						0,637
	$I_{1.3}$						0,471
C-P_3	$I_{3.1}$						0,466

Nota. Elaboración propia.

Es importante señalar que en el AFE hubo dos casos donde un ítem integró un campo de problema diferente al categorizado teóricamente, pero existe relación conceptual: i) el $I_{2.2c}$ representante de la S-$P_{2.5}$ sobre graficar la función de probabilidad de una variable aleatoria discreta, ha integrado el C-P_4 relativo a la función de distribución de una variable aleatoria discreta, sin embargo son dos funciones ligadas a aquella variable y conocida su función de probabilidad se puede determinar su función de distribución; ii) el $I_{3.1}$ representante de la S-$P_{3.1}$ sobre graficar la función de probabilidad de una variable aleatoria continua ha integrado el C-P_1 relativo a reconocer la variable aleatoria como una función, pero la función de probabilidad es un elemento fundamental de la variable aleatoria, tanto para las de tipo discretas como continuas.

También, en la estructura empírica del instrumento fueron descartados dos campos de problemas: i) C-P_3 vinculado a la función de probabilidad de una variable aleatoria continua, aunque parte de su contenido se incluyó en el C-P_1, mediante el $I_{3.1}$ representante de la S-$P_{3.1}$ sobre graficar aquella función; ii) C-P_8 relacionado con la aproximación de distribuciones de variables aleatorias discretas, aunque se abordó la aproximación de la binomial por la normal mediante el $I_{7.3}$ representante de la S-$P_{7.3}$ sobre describir la tendencia de datos representados gráficamente empleando aquella aproximación.

Finalmente, para comprobar si existen diferencias significas entre los distintos campos de problemas se realizó una matriz de correlaciones (ver Tabla 5.4). Se estudió la existencia de correlación entre las respuestas a las tareas para los distintos campos de problemas, debido a la naturaleza de la variable "tipo de respuesta" (cualitativa ordinal), fue realizado un análisis de correlaciones para escalas ordinales empleando el coeficiente no paramétrico de Tau-b de Kendall. Así se ha obtenido que, en dos contrastes de hipótesis en los que está involucrado el C-P_1 se rechaza la hipótesis nula "no existe correlación entre ambas variables" con un nivel de confianza del 95%. Por tanto, existe relación entre las respuestas al C-P_1 que involucra la variable aleatoria como función, con las respuestas al C-P_2 vinculado a la función de probabilidad de una variable aleatoria discreta, y con las respuestas al C-P_5 sobre algunos valores de posición central o de dispersión asociados a una variable aleatoria.

Tabla 5.4

Correlaciones entre los campos de problemas

	C-P_1	C-P_2	C-P_4	C-P_5	C-P_6	C-P_7
C-P_1		0,286*	0,144	0,355*	0,011	0,055
C-P_2			0,315*	0,054	0,353*	0,186
C-P_4				-0,032	0,167	-0,031
C-P_5					0,081	-0,039
C-P_6						0,248*
C-P_7						

Nota. Elaboración propia (* indica p-valor<0,05).

También se observa en la Tabla 5.4, que hay relación entre las respuestas al C-P_2 vinculado a la función de probabilidad de una variable aleatoria discreta y: i) el C-P_4 sobre la función de distribución de una variable aleatoria discreta; ii) el C-P_6 relativo a la distribución binomial. La menor correlación (0,248) se observa entre las respuestas a los campos de problemas sobre modelos probabilísticos, específicamente entre el C-P$_6$ vinculado a la distribución binomial y el C-P$_7$ sobre la distribución normal

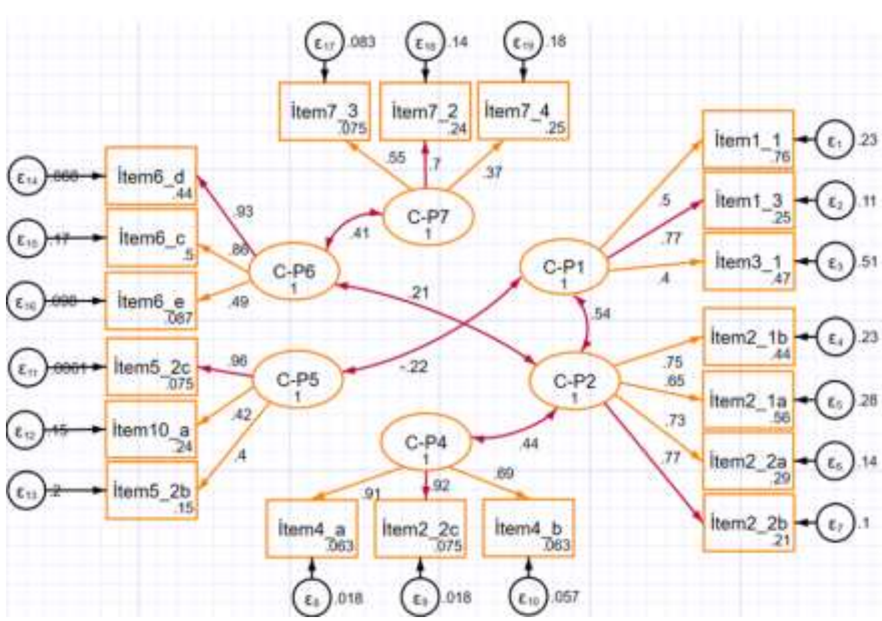
4.4.5.2 Análisis factorial confirmatorio

Para realizar el AFC se utilizó el modelo de ecuaciones estructurales (SEM) que según Escobedo et al. (2016) está compuesto por un conjunto de fases, las cuales se concretaron en tres: i) graficar la estructura teórica, espeficando las variables latentes (seis factores) y observadas (19 ítems); ii) estimar parámetros, como los coeficientes que conectan las variables latentes con las observadas (λ) y aquellos que relacionan las variables latentes con latentes (β); iii) evaluar el ajuste de la estructura teórica.

De esta manera, la Figura 5.3 expone la estructura del modelo de seis factores, que incluye los coeficientes de regresión (estandarizado) factor-ítem y factor-factor. La mayoría de los coefientes estandarizados factor-ítem varían entre 0,4 (I$_{3.1}$ en el factor C-P$_1$ y eI$_{5.2b}$ en el factor C-P$_5$) y 0,96 (I$_{5.2c}$ en el factor C-P$_5$), entonces según Galindo-Domínguez (2020) se consideran aceptables ($\lambda \geq 0{,}4$), salvo en el caso del I$_{7.4}$ con el factor C-P$_7$ cuyo coeficiente ha sido 0,37 por lo que se recomienda su revisión. Además, en cada factor la flecha roja representa el ítem con mayor carga factorial.

Figura 5.3

Estructura del modelo de seis factores



Nota. Elaboración propia.

Los resultados obtenidos sobre los coeficientes estandarizados factor-factor (ver Figura 5.3), en términos de Galindo-Domínguez (2020) indican que existe relación media ($|0{,}4| < \beta \leq |0{,}6|$), entre las respuestas al C-P_2 vinculado a la función de probabilidad de una variable aleatoria discreta, con las respuestas al C-P_1 que involucra la variable aleatoria como función ($\beta = 0{,}54$), y con las respuestas al C-P_4 en torno a la función de distribución de una variable aleatoria discreta ($\beta = 0{,}44$). Además hay relación media entre los campos de problemas sobre modelos probabilísticos, es decir, entre el C-P_6 distribución binomial y C-P_7 distribución normal ($\beta = 0{,}41$). Mientras que existe relación baja ($|0{,}2| < \beta \leq |0{,}4|$) entre las respuestas: i) al C-P_1 vinculado a la variable aleatoria como función y C-P_5 relativo a algunos valores de posición central o de dispersión asociados a la variable aleatoria ($\beta = -0{,}22$); ii) al C-P_2 sobre la función de probabilidad de la variable aleatoria discreta y C-P_6 relacionado con la distribución binomial ($\beta = 0{,}21$).

En relación con la evaluación del ajuste del modelo, la estructura empírica constituida por seis factores y 19 ítems tiene medidas de bondad de ajuste adecuadas: RMSEA=0,07; SRMR=0,08; TLI=0,87; y CFI=0,85. Sin embargo se advierte que las dos últimas medidas señaladas son discretamente inferiores a lo sugerido (Kline, 2005), según Peñacoba et al. (2022) aquello podría deberse al tamaño de la muestra inferior a 200.

La Tabla 5.5 expone los 19 ítems que componen la versión final del cuestionario y la fuente bibliográfica de cada uno de ellos.

Tabla 5.5

Ítems que componen versión final del cuestionario y fuente bibliográfica

Ítem	Fuente
1. De las situaciones que se describen a continuación: a) Lanzar dos dados para determinar el valor absoluto de la diferencia de los resultados. b) Tomar un artículo de un lote de 20 artículos de una fábrica, para determinar si está o no defectuoso. c) En la caída libre de un cuerpo, la relación entre el espacio recorrido y el tiempo desde su lanzamiento (donde la gravedad es $9{,}8\frac{\text{metros}}{\text{segundos}^2}$) d) Lanzar una moneda 4 veces y anotar el resultado que más se repite entre cara y sello. 1.1 Establece cuáles pueden ser variables aleatorias y descríbalas en palabras. 1.3 Enlista los valores de cada variable aleatoria descrita, además diseña un diagrama que relacione todos los posibles resultados de cada experimento aleatorio con estos valores.	1.1 Adaptado de Fernández et al. (2013) 1.3 Adaptado de Flores et al. (2014) y Salazar (2014)

2.1 Llamaremos lanzamiento a la acción de lanzar tres monedas al aire al mismo tiempo (de preferencia de la misma denominación). Ahora imagínate que se realizan 1000 lanzamientos y en cada uno de ellos se observa la variable “número de caras que ocurren”.
a) En la siguiente tabla anota en la fila superior, los posibles valores de la variable y en la fila inferior, el número de veces (o frecuencia) que crees que ocurra cada valor

Caras					Total
Frecuencia					

b) Anota la probabilidad que asignas a la ocurrencia de cada valor de la variable

Caras	0	1	2	3	Total
Probabilidad					

2.1 Extraído de Flores et al. (2014)

2.2 En un programa de televisión los concursantes hacen girar una flecha en una ruleta como en la figura y cada vez que la flecha se detiene en la zona de color rojo, se gana un premio y el concurso se trata de hacer girar la flecha 3 veces. En base a esta situación responde las siguientes preguntas:

a) ¿Cuál es la probabilidad de que un participante gane dos premios? Y ¿Cuál es la probabilidad de que un concursante no gane algún premio?
b) Escribe en la tabla siguiente la función de probabilidad de la variable aleatoria que relaciona el espacio muestral con el número de premios que gana un concursante.

X	0	1	2	3
f(x)				

c) En un gráfico represente la función de probabilidad anterior.

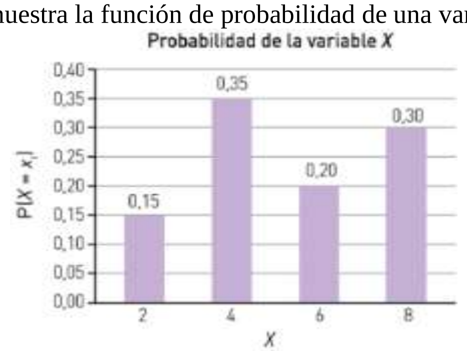
2.2a y 2.2b Extraídos de Salazar (2014)

2.2c Elaborado por autores

3.1 Considera la función de densidad f(x) = 0,2 definida en el intervalo [0, 5] y construye la gráfica de f

3.1 Extraído de Muñoz et al. (2013)

4. Para el gráfico que muestra la función de probabilidad de una variable aleatoria X:



a) Escribe su función de distribución.
b) Construye un gráfico que represente su función de distribución

4.a Elaborado por autores

4.b Extraído de Chacón et al. (2018)

5.2 Se define la variable aleatoria X: número de mascotas que tiene un estudiante. La función probabilidad asociada es:

x_i	0	1	2	3
$P(x_i)$	0,325	0,375	0, 225	0, 075

b) ¿Cuál es la varianza de la variable?
c) Si en el curso 3 medio de un liceo, el número medio de mascotas que tiene un estudiante es 1 y la desviación estándar es 0,95 ¿Cómo interpretarías cada uno de los resultados?

5.2b Extraído de Osorio et al. (2019)

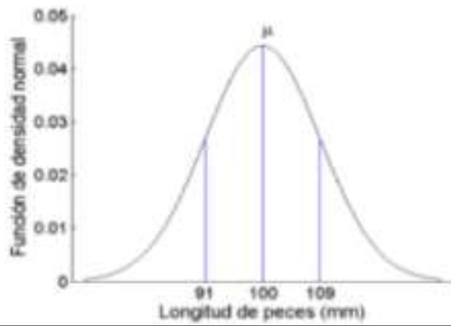
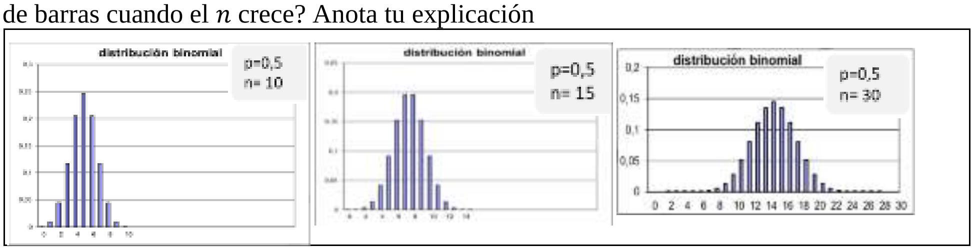
5.2c Elaborado por autores

6. Un profesor aplica a sus estudiantes un cuestionario sorpresa con cinco ítems de opción múltiple. Uno de los estudiantes no ha estudiado el material del cuestionario, y por tanto decide contestar los cinco ítems al azar, adivinando las respuestas sin leer las preguntas ni las opciones de respuestas.
Hoja de respuestas del cuestionario. Instrucciones: encierre en un círculo la mejor respuesta de cada ítem

1. A B C

6.c y 6.d Extraído de Alvarado y Retamal (2014)

6.e Elaborado por autores

2. A B C 3. A B C 4. A B C 5. A B C Antes de ver las respuestas correctas del cuestionario y encontrar cómo le fue a este estudiante, se considerarán algunos hechos que podrían suceder si un cuestionario se contesta de esta forma. c) ¿Cuál es la probabilidad de elegir las respuestas acertadas de los cinco ítems? d) ¿Cuál es la probabilidad de elegir las respuestas incorrectas de los cinco ítems? e) Calcula la media de las respuestas correctas.	
7.2 La longitud de cierto tipo de peces sigue una distribución normal con media $\mu = 100$ y varianza $\sigma^2 = 81$. ¿Cuál es la probabilidad de que uno de esos peces mida entre 82 y 91mm?  	7.2 Extraído de González y Ojeda (2017)
7.3 Observa las siguientes distribuciones binomiales (Figura 1) ¿Qué sucede con el gráfico de barras cuando el n crece? Anota tu explicación  *Figura 1.* Secuencia de distribuciones binomiales	7.3 Adaptado de MINEDUC (2019b)
7.4 V/F Justifica tu respuesta. "En una distribución normal, el 50% de las medidas caen por encima de la media"	7.4 Extraído de Tauber (2001)
10. Lee el siguiente titular, luego responda las preguntas propuestas  a) De manera resumida, explique cómo funciona la prueba PSU mencionando los datos representados en el gráfico.	10.a Elaborado por autores

Nota. Elaboración propia

4.4.5.3 Fiabilidad

Fue calculada el alfa de Cronbach del instrumento, cuyo valor obtenido (0,824) muestra que el instrumento posee una buena consistencia interna (George y Mallery, 2003). También se calculó aquel coeficiente de cada campo de problema (ver Tabla 5.6) y según George y Mallery (2003): el C-P_2 y C-P_4 poseen una fiabilidad buena (valores entre 0,8 y 09); el C-P_6 una fiabilidad aceptable (valores entre 0,7 y 0,8); y el C-P_1, C-P_5 y C-P_7 una fiabilidad pobre (valores entre 0,5 y 0,6).

Tabla 5.6

Coeficiente de alfa de Cronbach de cada campo de problema

C-P	Alfa de Cronbach
C-P_1	0,527
C-P_2	0,830
C-P_4	0,877
C-P_5	0,542
C-P_6	0,784
C-P_7	0,528

Nota. Elaboración propia.

4.4.6 Discusión y conclusiones

En el presente trabajo se analizó la validez de constructo y fiabilidad de un instrumento que valora la comprensión de variable aleatoria y sus distribuciones de probabilidad al término de la educación escolar. Respecto al primer aspecto, los resultados evidenciaron que la estructura empírica del cuestionario (19 ítems) está compuesta por seis factores que explican el 58% de la varianza total, cada uno vinculado a un campo de problema sobre: variable aleatoria (C-P_1), función de probabilidad de una variable aleatoria discreta (C-P_2), función de distribución de una variable aleatoria discreta (C-P_4), valores de posición central o de dispersión asociados a una variable aleatoria (C-P_5), distribución binomial (C-P_6) o distribución normal (C-P_7). Cada factor está integrado por tres a cuatro ítems, donde la mayoría de los coeficientes factor-ítem han sido aceptables según parámetros establecidos (Galindo-Domínguez, 2020) y el modelo propuesto ha presentado adecuadas medidas de ajuste de acuerdo con valores señalados en la literatura (Kline, 2005).

También el instrumento es fiable, ya que en términos de George y Mallery (2003) ha obtenido una buena consistencia interna. Por tanto, se puede concluir que el cuestionario compuesto por 19 ítems ha demostrado una validez de constructo y fiabilidad, entonces posee evidencias para evaluar articuladamente la comprensión de variable aleatoria y distribuciones binomial y normal en egresados de educación escolar chilena. De ahí la importancia del presente estudio, que propone un cuestionario con buenas características psicométricas y permite evaluar temas fundamentales de la educación estadística. Debido a que, en este contexto, las investigaciones sobre diseño de instrumentos, con robusta evidencia de fiabilidad y validez de constructo, abordan principalmente la valoración de actitudes hacia la estadística (Lavidas et al., 2020; Nasser, 2004; Persson, et al., 2019; Schulte, 2008), existiendo limitadas que evalúan comprensión o conocimiento de alguno de sus temas (Díaz, 2007; Ruiz 2019).

Además, el instrumento diseñado tiene diversas utilidades, tanto para evaluar la comprensión de estudiantes escolares posterior a un proceso de enseñanza-aprendizaje de variable aleatoria y/o distribuciones de probabilidad, como para diagnosticar la comprensión de estudiantes que acaban de ingresar a la universidad sobre dichos temas. Aún más, es una herramienta que posee información válida y fiable en torno a grandes ideas de la alfabetización probabilística, como la variable aleatoria y distribuciones binomial y normal (Jones et al., 2007), por lo que sería útil para el diseño de intervenciones educativas.

Sobre las limitaciones de este estudio, se señala el tamaño de la muestra debido a la Covid-19. Por lo que se proyecta la aplicación del cuestionario en una muestra de mayor tamaño y representatividad, para obtener información detallada en torno a la comprensión de variable aleatoria y sus distribuciones de probabilidad que poseen egresados de educación escolar chilena, especificando los conflictos semióticos que manifiesten en cada campo de problema.

Referencias

Alvarado, H. (2007). *Significado del teorema central del límite en la enseñanza de la estadística en ingeniería* [Tesis de Doctorado, Universidad de Granada].

Alvarado, H. y Retamal, M. (2014). Representaciones de la distribución de probabilidad binomial. En L. Andrade (Ed.), *Memorias del I Encuentro Colombiano de Educación Estocástica* (pp. 98-109). Acedest.

Alvarado, H., Estrella, S., Retamal, L. y Galindo, M. (2018). Intuiciones probabilísticas en estudiantes de ingeniería: implicaciones para la enseñanza de la probabilidad. *Revista Latinoamericana de Investigación en Matemática Educativa*, *21*(2), 131-156.

Amrani, H. y Zaki, M. (2015). Student’s conceptual difficulties with respect to the notion of random variable. *International Journal of Education, Learning and Development, 3*(9), 82-98.

Armah, G., Asiedu-Addo, S. y Owusu-Ansah, N. (2016). An investigation on the conceptual understanding of the standard deviation of fresh undergraduate students. *ASIO Journal of Chemistry, Physics, Mathematics and Applied Sciences*, *1*(3), 1-9.

Begué, N., Batanero, C., Gea, M. y Díaz-Levicoy, D. (2020). Razonamiento de estudiantes de bachillerato ante una situación binomial. *Tangram Revista de Educación Matemática, 3*(2), 27-50.

Bill, A., Watson, K. y Gayton, P. (2009). Guessing Answers to Pass a 5- item True False Test: Solving a Binomial Problem Three Different Ways. En R. Hunter, B. Bicknell y T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conferencia del Grupo de Investigación en Educación en Matemáticas de Australasia* (pp. 57-64). MERGA.

Bizet, V. y Ramos, E. (2022). Valoración de una situación didáctica para la enseñanza de variable aleatoria y distribución de probabilidad en la educación secundaria chilena. *Innovaciones Educativas, 24*(36), 21-36.

Bizet, V., Molina-Portillo, E., Ruz, F. y Contreras, J.M. (2022). What does the research tell us about the understanding of the random variables and its probability distribution*? Boletín de Estadística e Investigación Operativa*, *38*(3), 1-2.

Bizet, V., Molina-Portillo, E., Ruz, F. y Contreras, J.M. (2023a). Elaboración de una guía de situaciones-problema sobre variable aleatoria y sus aplicaciones a partir del currículo escolar chileno. *Educación Matemática*, *35*(1), 163-190.

Bizet, V., Molina-Portillo, E. y Contreras, J.M. (2023b). Situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad según libros de texto chilenos. *Profesorado, revista de currículum y formación de profesorado, 27*(2), 351-382.

Chacón, A., García, G., Rupín, P., Setz, J. y Villena, M. (2018). *Texto del estudiante de matemática 2° medio*. Ediciones SM.

Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología* [Tesis de Doctorado, Universidad de Granada].

Dinges, H. (2005). Variables, in particular random variables. En M. Hoffmann, J. Lenhard, y F. Seeger (Eds.), *Activity and sign grounding mathematics education* (pp. 305–311). Springer.

Del Pino, G. y Estrella, S. (2012). Educación estadística: relaciones con la matemática. *Pensamiento Educativo Revista de Investigación Educacional Latinoamericana*, *49*(1), 53-64.

Doukhan, C. y Gueudet, G. (2019). Students' difficulties at the secondary-tertiary transition: the case of random variables. En U. Jankvist, M. van den Heuvel-Panhuizen y M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 2464-2471). Freudenthal Group y Freudenthal Institute of the Utrecht University y ERME.

Escobedo, M., Hernández, J., Estebané, V. y Martínez G. (2016). Modelos de ecuaciones estructurales: características, fases, construcción, aplicación y resultados. *Ciencia y Trabajo*, *18*(55), 16-22.

Everitt, B. y Wykes, T. (2001). *Diccionario de estadística para psicólogos.* Ariel.

Fernández, F., Andrade, L. y Sarmiento, B. (2013). Rehaciendo el camino hacia la comprensión de la variable aleatoria. En P. Perry, C. Samper, Ó. Molina, L. Camargo, A. Echeverry, F. Fernández y B. Sarmiento (Eds.), *Aportes investigativos para el diseño curricular en geometría y estadística* (pp. 93–169). Universidad Pedagógica Nacional.

Flores, B., García, J. y Sánchez, E. (2014). Avances en la calidad de las respuestas a preguntas de probabilidad después de una actividad de aprendizaje con tecnología. En M. González, M. Codes, D. Arnau y T. Ortega (Eds.), *Investigación en Educación Matemática XVIII* (pp. 307-316). SEIEM.

Galindo-Domínguez, H. (2020). *Estadística para no estadísticos: una guía básica sobre la metodología cuantitativa de trabajos académicos.* 3Ciencias Editorial.

García, J., Medina, M. y Sánchez, E. (2014). Niveles de razonamiento de estudiantes de secundaria y bachillerato en una situación-problema de probabilidad. *Avance de Investigación en Educación Matemática, 6*, 5–23.

García, J. y Sánchez, E. (2013). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a una situación básica de variable aleatoria y distribución. En J. M. Contreras, G. R. Cañadas, M. M. Gea, y P. Arteaga (Eds.), *Actas de las jornadas virtuales de didáctica de la estadística*, *probabilidad y combinatoria* (pp. 417–424). Departamento de didáctica de la matemática de la Universidad de Granada.

García-García, J., Fernández, N., Arredondo, E. e Imilpán, I. (2022). The binomial distribution: historical origin and evolution of its problem situations. *Mathematics, 10*, 1-28.

Gea, M. (2014). *La correlación y regresión en bachillerato: análisis de libros de texto y del conocimiento de los futuros profesores* [Tesis de Doctorado, Universidad de Granada].

George, D. y Mallery, P. (2003). *SPSS for windows step by step: a simple guide and reference. 11.0 update.* Allyn and Bacon.

Godino, J. (1999). Implicaciones metodológicas de un enfoque semiótico- antropológico para la investigación en didáctica de las matemáticas. En T. Ortega (Ed.), *Investigación en Educación Matemática III* (pp. 196-212). SEIEM.

Godino, J. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, *22*(2.3), 237-284.

Godino, J. (2017). Construyendo un sistema modular e inclusivo de herramientas teóricas para la educación matemática. En J. M. Contreras, P. Arteaga, G. Cañadas, M. Gea, B. Giacomone y M. López-Martín (Eds.), *Actas del Segundo Congreso International Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos* (pp. 1-20). Departamento de Didáctica de la Matemática de la Universidad de Granada.

Godino, J., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM Mathematics Education*, *39*(1-2), 127-135.

Godino, J., Batanero, B. y Font, V. (2020). El enfoque ontosemiótico: implicaciones sobre el carácter prescriptivo de la didáctica. *Revista Chilena de Educación Matemática, 12*(2), 3-15.

González, Y. y Ojeda, A. (2017). Comprensión de la distribución normal en bachillerato. En L. Serna (Ed.), *Acta Latinoamericana de Matemática Educativa* (pp. 207-217). CLAME.

Guerrero, H., Ortiz, J. y Contreras, J.M. (2017). Evaluación del conocimiento sobre esperanza matemática y juegos equitativos en estudiantes de bachillerato. *Avance de Investigación en Educación Matemática, 11*, 107-125.

Hernández, R., Fernández, C. y Baptista, L. (2014). *Metodología de la investigación* (6ª ed.). McGraw Hill.

Jones, G., Langrall, C. y Mooney, E. (2007). Research in probability: Responding to classroom realities. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909-955). Information Age Publishing y NCTM.

Kline, R. (2005). *Principles and practice of structural equation modeling*. The Guilford Press.

Krippendorff, K. (1990). *Metodología de análisis de contenido: teoría y práctica*. Paidós.

Lavidas, K., Barkatsas, T., Manesis, D. y Gialamas, V. (2020). A structural equation model investigating the impact of tertiary students' attitudes toward statistics, perceived competence at mathematics, and engagement on statistics performance. *Statistics Education Research Journal, 19(*2), 27–41.

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, *30*(3), 1151-1169.

Martínez, M. (2020). La mejora de los logros escolares: la medida y valoración de la inspección educativa. *Educação e Pesquisa*, *46*, 1-17.

MINEDUC. Ministerio de Educación de Chile. (2015). *Bases Curriculares 7° básico a 2° medio. Gobierno de Chile.* Unidad de Currículum y Evaluación.

MINEDUC. Ministerio de Educación de Chile. (2019a). Chile. *Bases Curriculares 3º y 4º medio.* Unidad de Currículum y Evaluación.

MINEDUC. Ministerio de Educación de Chile. (2019b). *Programa de estudio matemática 4° medio para formación general.* Unidad de Currículum y Evaluación.

Muñoz, G., Gutiérrez, V. y Muñoz, S. (2013). *Texto matemática cuarto año medio.* Santillana.

Nasser, F. (2004). Structural Model of the Effects of Cognitive and Affective Factors on the Achievement of Arabic-Speaking Pre-service Teachers in Introductory Statistics. *Journal of Statistics Education*, *12*(1), 1-19.

NCTM. National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.

NGACBP. National Governors Association Center for Best Practices and CCSSO. Council of Chief State School Officers. (2010). *Common Core State Standars for Mathematics.* NGACBP and CCSSO.

Osorio, G., Norambuena, P., Romante, M., Gaete, D., Díaz, J., Celedón, J., Morales, K., Ortiz, N., Ramírez, P., Barrera, R. y Hurtado, Y. (2019). *Texto del estudiante de matemática 3° y 4° medio*. Ediciones SM.

Peñacoba, A., Muñoz, A.I. y Lafuente, M. B. (2022). Dimensiones latentes del proceso formativo del profesorado de Educación Secundaria: percepciones del alumnado. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, *25*(3), 221-237.

Persson, I., Kraus, K., Hansson, L. y Wallentin, F.Y. (2019). Confirming the structure of the survey of attitudes toward statistics(sats-36) by swedish students. *Statistics Education Research Journal*, *18*(1), 83–93.

Pfannkuch, M. (2018). International handbook of research in statistics education. En D. Ben-Zvi, K. Makar y J. Garfield (Eds.), *Reimagining curriculum approaches* (pp. 387-413). Springer.

Rodríguez-Santero, J. y Gil-Flores, J. (2019). Actitudes hacia la estadística en estudiantes de ciencias de la educación. Propiedades psicométricas de la versión española del Survey of Attitudes Toward Statistics (SATS-36). *Relieve, 25*(1),1-17.

Ruiz, B. (2013). *Análisis epistemológico de la variable aleatoria y comprensión de objetos matemáticos relacionados por estudiantes universitarios* [Tesis de Doctorado, Universidad de Granada].

Ruiz, K. (2019). *Comprensión del muestreo por alumnos chilenos de educación secundaria* [Tesis de Doctorado, Universidad de Granada].

Salazar, R. (2014). *La variable aleatoria con probabilidad desde la perspectiva de la teoría APOE* [Tesis de Máster, Pontificia Universidad Católica de Valparaíso].

Sánchez, E. y Carrasco, G. (2018). El razonamiento probabilístico de estudiantes de bachillerato en actividades de distribución binomial. En L. Rodríguez-Múniz, L. Múñiz-Rodríguez, Á. Aguilar-González, P. Alonso, F. García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 535- 543). SEIEM.

Sánchez, E., Carrasco, G. y Herrera, M. (2018). Fundamental ideas in the probabilistic reasoning of high - school students in binomial distribution activities. En M. Sorto, A. White, y L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics* (pp.1-6). IASE.

Sánchez, E. y Landín, P. (2014). Levels of probabilistic reasoning of high school students about binomial problems. En E. Chernoff y B. Sriraman (Eds.), *Probabilistic thinking presenting plural perspectives* (pp. 581-597). Springer.

Schulte, L. (2008). The development and validation of a teacher preparation program follow-up survey. *Journal of Statistics Education*, *16*(3), 1-10.

Shin, B. M. (2012). A Study on a Didactic Transposition Method and Students' Understanding for the Normal Distribution. *Journal of Educational Research in Mathematics, 22*(2), 117-136.

Stahl, S. (2006). The Evolution of the Normal Distribution. *Mathematics Magazine, 79*(2), 96-113.

Tauber, L. (2001). *La construcción del significado de la distribución normal a partir de actividades de análisis de datos* [Tesis de Doctorado, Universidad de Sevilla].

Valdez, J. y Salinas, J. (2019). Análisis de las respuestas de estudiantes de bachillerato a problemas sobre la distribución normal. En J. M. Contreras, M. Gea, M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística* (pp. 1-10). Departamento de didáctica de la matemática de la Universidad de Granada.

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D. y Verschaffel, L. (2003). The illusion of linearity: Expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics, 53*, 113–138.
Vilca, M. (2015). *Tipificación de los errores que se presentan al identificar una variable aleatoria de distribución binomial en problemas contextualizados* [Tesis de Máster, Pontificia Universidad Católica del Perú].
Williams, B., Onsman, A. y Brown, T. (2010). Exploratory factor analysis: a five-step guide for novices. *Australasian Journal of Paramedicine, 8*(3), 1-13.

4.5 Índice de dificultad e índice de discriminación de los ítems que componen la versión final del instrumento

Se calculó el índice de dificultad (Morales, 2009) de cada uno de los 19 ítems que componen la versión final del instrumento, aquel corresponde al porcentaje de aciertos resultante de los grupos con puntuaciones bajas (inferiores al percentil 25) y altas (superiores al percentil 75). Los valores muestran un índice bajo en cada ítem (ver Tabla 17), lo cual evidencia que ninguna de las situaciones-problemas sobre variable aleatoria y sus distribuciones de probabilidad fue fácil para los 80 participantes. En términos de Gómez (2014) 16 ítems fueron difíciles ($\text{ID} < 0{,}25$) y 3 presentaron una dificultad moderada ($0{,}25 \leq \text{ID} \leq 0{,}75$).

Tabla 17

Índices de dificultad y discriminación de los ítems

I	Índice de dificultad	Índice de discriminación	p-valor
1.1	0,09	0,2	**0,075**
1.3	0,09	0,2	0,007
3.1	0,23	0,5	0,002
2.1a	0,14	0,3	0,003
2.1b	0,27	0,42	0,014
2.2a	0,14	0,3	0,000
2.2b	0,14	0,3	0,000
2.2c	0,09	0,2	0,019
4.a	0,09	0,2	0,031
4.b	0,09	0,2	**0,113**
5.2b	0,18	0,22	0,047
5.2c	0	0	**0,112**
10.a	0	0	0,001
6.c	0,36	0,58	0,003
6.d	0,27	0,6	0,002
6.e	0,09	0,2	0,047
7.2	0,09	0,2	0,031
7.3	0,09	0,2	**0,113**
7.4	0,09	0,2	0,001

Nota. Elaboración propia (Resultados estadísticamente significativos a un nivel de 0,05).

También fue calculado el índice de discriminación (Morales, 2009) de cada ítem, entendido como la diferencia entre el porcentaje de aciertos en el grupo de puntuaciones altas y el porcentaje de lo mismo en el grupo con puntuaciones bajas. Estos valores (ver Tabla 17) evidencian según Ebel y Frisbie (1991) que: cuatro ítems presentaron una discriminación muy buena ($\mathrm{IDC} \geq 0{,}40$), tres una condición adecuada ($0{,}30 \leq \mathrm{IDC} \leq 0{,}39$), 10 índice regular ($0{,}20 \leq \mathrm{IDC} \leq 0{,}29$), dos tuvieron una fuerza de discriminación pobre ($IDC < 0{,}20$). Además, fue realizado un contraste de hipótesis para muestras independientes, mediante la prueba de U de Mann-Whitney, para esta se consideró dos grupos, los de puntuaciones bajas y altas. Así en 15 ítems se rechazó la hipótesis nula (p-valor<0,05) y concluyó que existen diferencias significativas entre las respuestas correctas de los grupos de alto y bajo rendimiento, es decir, en la mayoría de los ítems hay una adecuada discriminación.

Debido a que la fuerza de discriminación de un ítem está relacionada con su grado de dificultad en función del contenido evaluado, fueron analizados estos índices en conjunto para los casos conflictivos (dos ítems). Los reactivos $I_{5.2c}$ y $I_{10.a}$ poseen menor índice de difícil y discriminación, puntualmente la mayoría de los estudiantes respondió erróneamente, tanto los de bajo como alto rendimiento, por lo que requieren una revisión. Para no perder contenido valorado por cada ítem señalado se propuso mejorar su redacción como se aprecia en la Figura 24.

Figura 24

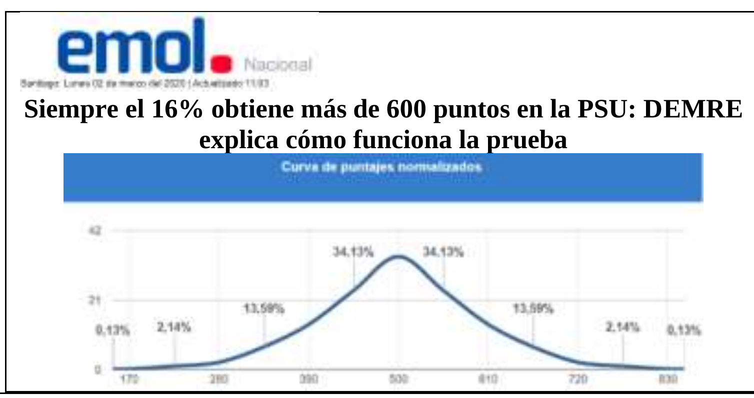
Modificación en la redacción del ítem 5.2c e ítem 10.a

5.2

c) Si en el curso 3 medio de un liceo, el número medio de mascotas que tiene un estudiante es 1 y la desviación estándar es 0,95 ¿Cómo interpretarías cada uno de los resultados? (Inicial)

c) En el curso 3 medio de un liceo, el número medio de mascotas que tiene un estudiante es 1 y la desviación estándar es 0,95 ¿Cómo interpretarías cada medida? (Modificado)

10. Lee el siguiente titular, luego responda la pregunta propuesta



a) De manera resumida, explique cómo funciona la prueba psu mencionando los datos representados en el gráfico (Inicial) a) Explique cómo funciona la prueba psu solamente mencionando los datos representados en el gráfico (Modificado)

Nota. Elaboración propia.

4.6 Significado institucional evaluado en el instrumento sobre variable aleatoria y distribuciones binomial y normal

Las Tablas 18 y 19 presentan el significado institucional de variable aleatoria evaluado en el cuestionario por 13 ítems: ítem 1 ($I_{1.1}$ e $I_{1.3}$), ítem 2.1 ($I_{2.1a}$ e $I_{2.1b}$), ítem 2.2 ($I_{2.2a}$, $I_{2.2b}$ e $I_{2.2c}$), ítem 3.1 ($I_{3.1}$), ítem 4 ($I_{4.a}$ e $I_{4.b}$), ítem 5 ($I_{5.2b}$ e $I_{5.2c}$) e ítem 10.a ($I_{10.a}$). En tanto que las Tablas 20 y 21 exponen el significado institucional de distribuciones binomial y normal valorado en el cuestionario por 6 ítems: ítem 6 ($I_{6.c}$, $I_{6.d}$ e $I_{6.e}$) e ítem 7 ($I_{7.2}$, $I_{7.3}$ e $I_{7.4}$). Estos significados institucionales sintetizan el análisis a priori realizado a los 19 ítems (ver apartado 4.3), donde fueron identificados los objetos matemáticos primarios involucrados en las prácticas desarrolladas para resolver cada uno.

El significado institucional de variable aleatoria evaluado en el cuestionario (ver Tablas 18 y 19) en contraste con su significado institucional de referencia según el currículo escolar chileno (ver Anexo 1), aborda la mayoría de las situaciones-problemas propuestas en este último (16 de 21), exceptuando (Bizet et al., 2023c, p.180): la S-P$_{1.3}$ identificar dominio de una variable aleatoria finita, la S-P$_{3.2}$ calcular algunas probabilidades asociadas a una variable aleatoria continua, la S-P$_{3.3}$ determinar si la función dada es una función de probabilidad de una variable aleatoria continua, la S-P$_{5.1}$ calcular la media o valor esperado de una variable aleatoria discreta, y la S-P$_{5.4}$ calcular la desviación estándar de una variable aleatoria discreta. También abarca todos los tipos de lenguaje incluidos en el significado institucional de referencia.

Tabla 18

Situaciones-problemas sobre variable aleatoria evaluadas en la versión final del instrumento

Situaciones-problemas	Ítems						
	1	2.1	2.2	3.1	4	5	10.a
S-P$_{1.1}$ Diferenciar entre variables aleatorias y variables con dependencia funcional	X						
S-P$_{1.2}$ Definir variables aleatorias finitas en experimentos aleatorios	X						
S-P$_{1.4}$ Identificar recorrido o valores de una variable aleatoria finita	X						
S-P$_{3.1}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua				X			
S-P$_{2.1}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial		X					

S-$P_{2.2}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico		X			
S-$P_{2.3}$ Definir la función de probabilidad de una variable aleatoria utilizando la terminología $P(X = x_i)$		X			
S-$P_{2.4}$ Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta	X				
S-$P_{2.5}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta		X			
S-$P_{4.1}$ Determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta			X		
S-$P_{4.2}$ Definir la función de distribución de una variable aleatoria discreta			X		
S-$P_{4.3}$ Representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta			X		
S-$P_{5.2}$ Interpretar la media o valor esperado de una variable aleatoria discreta				X	
S-$P_{5.3}$ Calcular la varianza de una variable aleatoria discreta				X	
S-$P_{5.5}$ Interpretar la desviación estándar de una variable aleatoria discreta				X	
S-$P_{5.6}$ Identificar la media y desviación estándar de una variable aleatoria continua					X

Nota. Extraída de Bizet et al. (2023c).

Sobre los conceptos, aborda todos los propuesto por el currículo escolar chileno, focalizándose en la probabilidad y estadística, aunque incorpora algunos conceptos del eje temático de álgebra y funciones de los lineamientos curriculares: i) función, su dominio y su recorrido, y ii) variable dependiente y variable independiente.

En relación con las proposiciones, incluye las consideradas en el significado institucional de referencia, además de otras ligadas al eje de probabilidad y estadística del currículo escolar de Chile, como por ejemplo la regla del producto y la regla de Laplace. Por aquel motivo se enriquece el conjunto de procedimientos abordados en el cuestionario. Respecto a los argumentos, principalmente se centran en la representación gráfica para justificar respuestas, el cual ha sido considerado en el significado institucional de referencia.

Tabla 19

Lenguajes, conceptos, proposiciones, procedimientos y argumentos sobre variable aleatoria evaluadas en la versión final del instrumento

Tipo de objeto matemático	Objeto matemático	Ítems						
		1	2.1	2.2	3.1	4	5	10.a
Lenguajes	Verbal	X	X	X	X	X	X	X
	Numérico		X	X		X	X	
	Simbólico	X	X	X	X	X	X	X
	Gráfico	X		X	X	X		X
	Tabular		X	X			X	
Conceptos	Experimento aleatorio	X	X	X				
	Espacio muestral	X	X	X				
	Suceso aleatorio	X	X	X			X	
	Sucesos independientes			X				

	Función (dominio y recorrido)	X			X	X		
	Variable dependiente e independiente	X						
	Variable aleatoria	X	X	X	X	X	X	X
	Conjunto finito	X	X	X				
	Enfoque frecuencial de probabilidad		X					
	Enfoque clásico de probabilidad		X	X				
	Distribución de frecuencias		X					
	Función de probabilidad		X	X		X	X	
	Función de distribución					X		
	Intervalo				X			X
	Función de densidad				X			
	Combinatoria			X				
	Media						X	X
	Varianza						X	
	Desviación estándar						X	X
	Distribución binomial			X				
	Función de probabilidad de la binomial			X				
	Distribución normal							X
	Función de densidad de la normal							X
Proposiciones	Caracterización de la variable aleatoria y variables con dependencia funcional	X						
	Caracterización de variables aleatorias discreta y continua	X	X	X	X	X	X	X
	Convergencia, al crecer el número de ensayos la frecuencia relativa se va estabilizando		X					
	Regla de Laplace		X	X				
	Caracterización de la variable aleatoria mediante su función de probabilidad		X	X		X	X	
	La media es un representante de los datos a partir de los que ha sido calculada y no tiene por qué ser igual a uno de los valores de los datos						X	
	Propiedades de la función de probabilidad		X	X		X		
	Regla del producto			X				
	La probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento			X				
	Caracterización de la variable aleatoria mediante su función de distribución					X		
	Caracterización de la distribución binomial			X				
	Propiedades de la función de densidad				X			
	Propiedad de los intervalos centrales ($\pm 3\sigma$)							X
Procedimientos	Construcción del espacio muestral	X						
	Partición disjunta del espacio muestral	X	X	X				
	Construcción de un diagrama sagital	X						
	Reproducción de experimentos aleatorio conservando las condiciones iniciales		X					
	Cálculo de frecuencias (relativas o absolutas)		X					
	Cálculo de probabilidades con la regla de Laplace		X	X				
	Cálculo de combinaciones			X				
	Cálculo de probabilidad compuesta con la regla del producto			X				
	Construcción de un gráfico de barras			X				
	Lectura e interpretación de un gráfico de barras					X		
	Construcción de la gráfica de una función real en el plano cartesiano				X	X		

	Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de distribución					X		
	Lectura de una tabla de probabilidad						X	
	Cálculo de medida de dispersión asociada a una variable aleatoria discreta empleando su expresión algebraica						X	
	Lectura e interpretación de una curva normal							X
	Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial			X				
Argumentos	Verbal-deductivo	X	X	X			X	
	Mediante representación gráfica	X		X	X	X		X

Nota. Elaboración propia.

El significado institucional de distribuciones binomial y normal evaluado en el cuestionario (ver Tablas 20 y 21) en contraste con su significado institucional de referencia según el currículo escolar chileno (ver Anexo 1), incluye gran parte de las situaciones-problemas propuestas en este último (7 de 12), exceptuando (Bizet et al., 2023c, p.181-182): la S-$P_{6.5}$ determinar probabilidades asociadas a una distribución binomial empleando una herramienta tecnológica, la S-$P_{7.1}$ identificar situaciones que pueden modelizarse a través de una distribución normal, la S-$P_{7.5}$ determinar probabilidades asociadas a una distribución normal empleando una herramienta tecnológica, la S-$P_{8.1}$ calcular los parámetros asociados a una distribución normal, y la S-$P_{8.2}$ calcular probabilidades en el contexto de aproximación de la distribución binomial a la normal. Además, aborda la mayoría de los tipos de lenguajes incluidos en los lineamientos curriculares chilenos, a excepción del lenguaje gráfico en el contexto de la distribución binomial.

Tabla 20

Situaciones-problemas sobre distribuciones binomial y normal evaluadas en la versión final del instrumento

Situaciones-problemas	Ítems			
	6	7.2	7.3	7.4
S-$P_{6.1}$ Identificar situaciones que pueden modelizarse a través de una distribución binomial	X			
S-$P_{6.2}$ Determinar los parámetros asociados a una distribución binomial como n (número de ensayos), p (probabilidad de éxito) y q (probabilidad de fracaso)	X			
S-$P_{6.3}$ Calcular la media de una distribución binomial	X			
S-$P_{6.4}$ Determinar probabilidades asociadas a una distribución binomial de forma manual	X			
S-$P_{7.2}$ Determinar probabilidades asociadas a una distribución normal de forma manual		X		
S-$P_{7.3}$ Describir la tendencia de datos representados gráficamente empleando la aproximación de la binomial por la normal			X	
S-$P_{7.4}$ Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal				X

Nota. Extraída de Bizet et al. (2023c).

Respecto a los conceptos, abarca la mayoría de los propuesto en el significado institucional de referencia, aunque excluye en el contexto de la distribución binomial su función de distribución, su varianza y su desviación estándar, y en el ámbito de la normal su moda. También se centra en los conceptos del eje temático estadística y probabilidad de los lineamientos curriculares chilenos, aunque incorpora del eje álgebra y funciones el concepto de intervalos.

Sobre las proposiciones, aborda las involucradas en el significado institucional de referencia. Además, en el contexto de la binomial incorpora otras proposiciones vinculadas al eje de probabilidad y estadística del currículo chileno: i) aquella que indica que la probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento, ii) la regla del producto, y iii) la regla de Laplace. Debido a dicho motivo se enriquece el conjunto de procedimientos relativos a la distribución binomial valorados en el cuestionario. Con relación a los procedimientos en torno a la distribución normal, incluye los considerados en el significado institucional de referencia, aún más adquieren relevancia los procedimientos de verificar sus propiedades o sus características y leer e interpretar una curva normal y un gráfico de barras.

En relación con los argumentos, se focalizan en el razonamiento verbal-deductivo y la representación gráfica para validar soluciones, aunque sólo este último ha sido incluido en el significado institucional de referencia de las distribuciones binomial y normal.

Tabla 21

Lenguajes, conceptos, proposiciones, procedimientos y argumentos sobre distribuciones binomial y normal evaluadas en la versión final del instrumento

Tipo de objeto matemático	Objeto matemático	Ítems 6	7.2	7.3	7.4
Lenguajes	Verbal	X	X	X	X
	Numérico	X	X		
	Simbólico	X	X	X	X
	Gráfico		X	X	
	Tabular		X		
Conceptos	Experimento aleatorio	X		X	
	Espacio muestral	X			
	Suceso aleatorio	X			
	Sucesos independientes	X			
	Variable aleatoria	X	X	X	
	Enfoque clásico de probabilidad	X	X		
	Intervalo		X		
	Combinatoria	X			
	Media	X	X	X	X
	Mediana				X
	Varianza		X		

	Desviación estándar		X	X	
	Distribución binomial	X		X	
	Función de probabilidad de la binomial	X			
	Distribución normal		X	X	X
	Simetría		X		X
	Función de densidad de la normal		X		X
	Distribución normal estándar		X		
Proposiciones	Caracterización de variables aleatorias discreta y continua	X	X	X	
	Regla de Laplace	X			
	Propiedades de la función de probabilidad	X			
	Regla del producto	X			
	La probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento	X			
	Caracterización de la distribución binomial	X			
	la media es un representante de los datos a partir de los que ha sido calculada y no tiene por qué ser igual a uno de los valores de los datos	X			
	Caracterización de la distribución normal		X	X	
	Propiedades de la función de densidad de la normal				X
	Propiedad intervalos centrales ($\pm 3\sigma$)		X		
	Propiedades para el cálculo de probabilidades con distribución normal estándar		X		
	Condición para aproximar una distribución binomial a una normal			X	
Procedimientos	Cálculo de probabilidad (de éxito) con la regla de Laplace	X			
	Cálculo de combinaciones	X			
	Cálculo de probabilidad compuesta con la regla del producto	X			
	Lectura e interpretación de un gráfico de barras			X	
	Lectura e interpretación de una curva normal				
	Cálculo de probabilidad empleando la propiedad $\pm 3\sigma$		X		
	Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial	X			
	Cálculo de media de una binomial mediante su expresión algebraica	X			
	Verificación de propiedades o característica de la distribución normal			X	X
	Tipificación		X		
	Cálculo de probabilidades utilizando la función de densidad de la normal estándar		X		
	Lectura de tabla de la función de densidad de la distribución normal estándar		X		
	Verificación de condiciones para la aproximación a una distribución normal			X	
Argumentos	Verbal-deductivo	X	X		X
	Mediante representación gráfica			X	

Nota. Elaboración propia.

Por último, ciertos tipos de argumentos del significado institucional de referencia sobre variable aleatoria y distribuciones de probabilidad han sido excluidos del instrumento de evaluación. Por ejemplo, justificación mediante simulación con herramienta tecnológica u objetos manipulables, debido a que no se ajustan al foco de la presente investigación.

CAPÍTULO 5: EVALUACIÓN DE LA COMPRENSIÓN

5.1 Introducción

El presente capítulo expone el Estudio 6. Comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar chilena, donde aquella comprensión se evalúa en 101 estudiantes universitarios mediante la aplicación de un cuestionario previamente validado (Capítulo 4). Primero se expone la metodología utilizada, incluyendo la descripción de la muestra y los procedimientos desarrollos para el análisis de los datos.

A continuación, se dan a conocer los resultados conseguidos del análisis detallado de cada uno de los ítem, donde se describen las categorías (y subcategorías) de respuestas, expone un ejemplo de cada una e indagan los porcentajes de respuestas correctas, parciales, en banco y erróneas. Después, son descritos los conflictos semióticos principalmente detectados en las respuestas parcialmente correctas y erróneas de los participantes.

Posteriormente se presenta el significado personal logrado por los participantes en torno a la variable aleatoria y distribuciones de probabilidad. Puntualmente se identifican los objetos matemáticos primarios que los estudiantes chilenos lograron emplear para solucionar las tareas propuestas o avanzar parcialmente en su solución, con su respectivo porcentaje. También son reconociendo aquellos conceptos, lenguajes, procedimientos, proposiciones y argumentos que los estudiantes evidenciaron no manipular en la resolución de las situaciones-problemas evaluadas en el cuestionario.

Luego, se exponen una síntesis de los resultados, incluyendo algunos estadísticos descriptivos de la puntuación conseguida por los participantes en cada ítem, en cada campo de problema y en el cuestionario, también es analizada esta información según los grupos que integran la muestra. Para finalizar, se presenta la discusión de los resultados obtenidos y las conclusiones del estudio de evaluación.

5.2 Estudio 6. Comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar chilena

En consideración con los antecedentes, exhibidos en el Capítulo 2, y el análisis al tratamiento de los temas de interés en el currículo chileno, expuesto en el Capítulo 3, surgió la necesidad de realizar el presente Estudio 6. Este tiene por objetivo evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en estudiantes egresados de educación escolar chilena, fundamentado en el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos (Godino et al., 2020).

Con el cuestionario utilizado, cuyo proceso de contrucción y validez (de contenido y constructucto) fue expuesto en el Capítulo 4, se busca establecer una aproximación al significado personal logrado por los participantes en torno a variable aleatoria y distribuciones binomial y normal, y analizar la correspondencia con el significado institucional evaluado en el cuestionario.

5.2.1 Metodología

El estudio fue abordado desde un enfoque mixto (Hernández et al., 2014), debido a que incluyó tanto un componente cualitativo, pues se definieron y describieron categorías de respuestas en cada ítem, como un componente cuantitativo, ya que se realizó un análisis estadístico descriptivo para cada uno de los ítems.

5.2.1.1 Descripción de la muestra

La población de interés de la investigación fue estudiantes egresados de educación escolar chilena, ya que el cuestionario diseñado principalmente evalúa temas que son enseñados en las escuelas chilenas desde 2° medio (15 a 16 años) hasta 4° medio (17 a 18 años). Debido a la pandemia de la Covid-19, fue complejo acceder a estudiantes de primer año de universidad, por lo que se utilizó una muestra dirigida (no probabilística) de 101 estudiantes universitarios chilenos, quienes no habían cursado en la universidad alguna asignatura sobre tópicos estocásticos.

Los participantes fueron organizados en dos grupos según la carrera universitaria que estudiaban (ver Tabla 22). El primer grupo (G_1) estuvo compuesto por 56 estudiantes (30 mujeres y 26 hombres, con edades entre 17 y 23 años) de la carrera de Pedagogía en Matemática: 12 de primer año, 30 de segundo y 14 de tercero. El segundo grupo (G_2) integrado por 45 estudiantes (42 mujeres y 3 hombres, con edades entre 17 y 23 años) de primer año de la carrera de Pedagogía en Educación Especial.

Tabla 22

Distribución de la muestra por carrera universitaria y género

Edad	Pedagogía en Matemática		Pedagogía en Educación Especial		Total
	Mujer	Hombre	Mujer	Hombre	
17 a 18 años	4	3	16	3	26
19 a 20 años	20	6	19	0	45
21 a 22 años	3	7	5	0	15
23 años o más	3	10	2	0	15
Total	30	26	42	3	101

Nota. Elaboración propia.

Para asegurar que los participantes estudiaron en la escuela temas sobre estocástica y específicamente aquellos evaluados en el cuestionario, al inicio del instrumento se plantearon dos preguntas cerradas en torno a la asignatura de matemáticas, cada una con cuatro opciones de respuestas (ver Tablas 23 y 24): 1) ¿usted estudió temas sobre estadística y probabilidad durante la educación media?; y 2) ¿usted estudió los temas de variable aleatoria, distribución binomial y distribución normal durante la educación media?

Tabla 23

Frecuencia y porcentaje de respuestas en la pregunta 1

Opciones de respuesta	Pedagogía en Matemática	Pedagogía en Educación Especial	Total
No	6(10,7%)	0	6(5,9%)
No recuerdo	5(8,9%)	5(11,1%)	10(9,9%)
Si, algunos años	31(55,4%)	32(71,1%)	63(62,4%)
Si, cada año	14(25%)	8(17,8%)	22(21,8%)
Total	56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

En la asignatura de matemática durante la educación media, sólo el 5,9% de la muestra afirmó no haber estudiado temas sobre estadística y probabilidad, y el 9,9% no recordó si estudió o no aquellos (ver Tabla 23). Además, como expone la Tabla 24, alrededor del 13% de los participantes señaló que estudió la variable aleatoria y distribuciones binomial y normal, y el 41,6% reconoció haber estudiado algunos de estos temas.

Tabla 24

Frecuencia y porcentaje de respuestas en la pregunta 2

Opciones de respuesta	Pedagogía en Matemática	Pedagogía en Educación Especial	Total
No	15(26,8%)	4(8,9%)	19(18,8)
No recuerdo	11(19,6%)	16(35,6%)	27(26,7%)
Si, algunos temas	24(42,9%)	18(40%)	42(41,6)
Si, todos los temas	6(10,7%)	7(15,6%)	13(12,9%)

Nota. Elaboración propia.

El instrumento fue aplicado entre los meses de octubre y noviembre de 2021, bajo la modalidad en línea por motivo de la Covid-19. Los participantes respondieron el cuestionario voluntaria e individualmente en una sesión de sus asignaturas universitarias, con la presencia de su profesor e investigador, mediante un link enviado a su correo personal y en un tiempo de 90 minutos.

5.2.1.2 Prodecimientos de análisis de datos

Para el análisis de las respuestas proporcionadas por los participantes se establecieron categorías en cada ítem, empleando la técnica de análisis de contenido (Krippendorff, 1990) y mediante un proceso cíclico e inductivo. Las categorías han sido definidas tanto a partir del análisis a priori realizado a los ítems (ver Capítulo 4), como desde la indagación a las respuestas proporcionas por los participantes en cada uno de ellos. Además, a estas se les asignó un símbolo que posee un significado y un valor numérico (para facilitar el análisis cuantitativo): respuesta correcta se le otorgó el símbolo C_2 y el valor 2; respuesta parcialmente correcta se le designó el símbolo C_1 y valor 1; respuesta en blanco se le asignó el símbolo C_0 y valor 0; y respuesta errónea se le concedió el símbolo E y el valor 0. También según el ítem de estudio fueron definidas subcategorías en las respuestas correctas ($C_{2.n}$), parcialmente correctas ($C_{1.n}$) y erróneas (E_n).

Luego en las respuestas correctas y parcialmente correctas de cada ítem, han sido reconocidos los objetos matemáticos primarios movilizados por los estudiantes para su solución y su frecuencia. Además fueron identificados conflictos semióticos manifestados por los participantes, principalmente en las respuestas a los ítems clasificadas como parcialmente correctas e incorrectas.

El análisis estadístico descriptivo de los datos fue realizado mediante el programa SPSS Statistics (versión 25). Primero se presentaron en tablas la distribución de frecuencias de las categorías y subcategorías de respuestas observada en cada ítem. Después se calcularon algunos estadísticos descriptivos de la puntuación obtenida por los participantes en cada ítem, en cada campo de problema y en el cuestionario, también se llevaron a cabo aquellos cálculos según los grupos que integran la muestra. Específicamente se han determinado los estadísticos descriptivo como la media, desviación estándar, mínimo y máximo, con el propósito de entregar medidas cuantitativas de la comprensión de los participantes sobre variable aleatoria y distribuciones binomial y normal.

5.2.2 Resultado por ítem

5.2.2.1. Resultados en el ítem 1

El primer ítem expuesto en la Figura 25 está integrado por: el $I_{1.1}$ adaptado de la secuencia de instrucción de Fernández et al. (2013) que valora la S-$P_{1.1}$ diferenciar entre variables aleatorias y variables con dependencia funcional y la S-$P_{1.2}$ definir variables aleatorias finitas en experimentos aleatorios (Bizet et al., 2023c, p.187); y el $I_{1.3}$ adaptado de Flores at al. (2014) y Salazar (2014) que evalúa la S-$P_{1.4}$ identificar recorrido o valores de una variable aleatoria finita (Bizet et al., 2023c, p.187).

Figura 25

Ítem 1 del cuestionario

1. De las situaciones que se describen a continuación: a) Lanzar dos dados para determinar el valor absoluto de la diferencia de los resultados. b) Tomar un artículo de un lote de 20 artículos de una fábrica, para determinar si está o no defectuoso. c) En la caída libre de un cuerpo, la relación entre el espacio recorrido y el tiempo desde su lanzamiento (donde la gravedad es $9{,}8\frac{\text{metros}}{\text{segundos}^2}$) d) Lanzar una moneda 4 veces y anotar el resultado que más se repite entre cara y sello. 1.1 Establece cuáles pueden ser variables aleatorias y descríbalas en palabras. 1.3 Enlista los valores de cada variable aleatoria descrita, además diseña un diagrama que relacione todos los posibles resultados de cada experimento aleatorio con estos valores.

Nota. Adaptado de Fernández et al. (2013), Flores at al. (2014) y Salazar (2014).

En el $I_{1.1,}$ el análisis de las respuestas de los participantes permitió obtener las categorías y frecuencias expuestas en la Tabla 25.

Tabla 25

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{1.1}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	3(5,36%)	0	3(2,97%)
	$C_{2.2}$	2 (3,57%)	0	2 (1,98%)
Respuesta parcialmente correcta	$C_{1.1}$	23(41,07%)	15(32,33%)	38(37,62%)
	$C_{1.2}$	15(26,79%)	2(4,44%)	17(16,83%)
Respuesta en blanco	No aplica	6(10,71%)	15(33,33%)	21(20,79%)
Respuesta errónea	E_1	1(1,79%)	3(6,67%)	4(3,96%)
	E_2	3(5,36%)	3(6,67%)	6(5,94%)
	E_3	3(5,36%)	7(15,56%)	10(9,9%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como se aprecia en la Tabla 25, alrededor del 5% de la muestra respondió correctamente en el $I_{1.1}$. Solamente el 1,98% de los estudiantes lograron identificar las dos situciones propuestas (a y b) que pueden ser variables aleatorias y las describieron en lenguaje verbal ($C_{2.2}$): a) valor absoluto de la diferencia entre los números obtenidos al lanzar dos dados; y b) representa el estado del artículo en un determinado momento, siendo 1 si el artículo está defectuoso y 0 en caso contrario. Además el 2,97% de los participantes reconocieron las dos situaciones, pero solamente definieron una variable aleatoria ($C_{2.1}$). Ejemplos de esta categoría de respuesta se aprecian en la Figura 26.

Figura 26

Ejemplos de respuestas correctas en el $I_{1.1}$

Respuesta del estudiante 23 clasificada en la $C_{2.2}$

a. La variable aleatoria es el valor absoluto de la diferencia de los dos valores que salen al lanzar los dados
b. Se puede definir la variable aleatoria como 1 si está defectuoso o 0 si no lo está

Respuesta del estudiante 16 clasificada en la $C_{2.1}$

a) el resultado de cada dado es una variable aleatoria, asumiendo que los dados no están cargados, por ende la diferencia de los resultados será aleatoria con ciertas probabilidades
b) la eleccion del articulo puede ser aleatoria

Nota. Elaboración propia.

También sobre el 50% de la muestra respondió parcialmente correcto al $I_{1.1}$, pues el 16,83% de los estudiantes identificaron a lo más las dos situaciones propuestas que pueden ser variables aleatorias sin definirlas ($C_{1.2}$) y el 37,62% de los participantes, además de aquello, reconocieron incorrectamente otra situación ($C_{1.1}$). La Figura 27 ejemplifica estos tipos de respuesta.

Figura 27

Ejemplos de respuestas parcialmente correctas en el $I_{1.1}$

Respuesta del estudiante 14 clasificada en la $C_{1.2}$

la a) puede ser variable aleatoria porque al lanzar dos dados se obtienen diferentes combinaciones, los cuales después se les puede asignar un valor de probabilidad

Respuesta del estudiante 11 clasificada en la $C_{1.1}$

a)Lanzar dos dados para determinar el valor absoluto de la diferencia de los resultados
b)Tomar un articulo de un lote de 20 articulos de una fabrica, para determinar si esta o no defectuoso.
d)Lanzar una moneda 4 veces y anotar el resultado que mas se repite.
No sabría describirlo en palabras.

Nota. Elaboración propia.

Sin embargo, alrededor del 20% de la muestra no respondió o dijo no saber sobre lo preguntado en el $I_{1.1}$ (C_0). Aún más, en dicho ítem el 19,8% de los estudiantes respondieron de manera incorrecta debido a que: el 3,96% de los participantes confundieron una variable aleatoria con un experimento aleatorio (E_1); el 5,94% de los jóvenes identificaron erróneamente una situación (c o d) que no puede ser variable aleatoria (E_2); y el otro 9,9% propusieron incorrectamente dos situaciones que no pueden ser variables aleatorias (c y d), además de reconocer bien una situación (a o b) (E_3). Ejemplos de aquellos errores se muestran en la Figura 28.

Figura 28

Ejemplos de respuestas erróneas en el $I_{1.1}$

Respuesta del estudiante 64 clasificada en E_1	Respuesta del estudiante 45 clasificada en E_2	Respuesta del estudiante 60 clasificada en E_3
a) lanzar dos dados d) lanzar una moneda 4 veces	La respuesta es D	a, c, d

Nota. Elaboración propia.

Por otra parte, la indagación de las respuestas al $I_{1.3}$ proporcionada por los estudiantes hizo posible establer las categorías y frecuencias presentadas en la Tabla 26.

Tabla 26

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{1.3}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	2(3,57%)	0	2(1,98%)
	$C_{2.2}$	1(1,79%)	0	1(0,99%)
Respuesta parcialmente	$C_{1.1}$	6(10,71%)	2(4,44%)	8(7,92%)
correcta	$C_{1.2}$	6(10,71%)	0	6(5,94%)
Respuesta en blanco	No aplica	27(48,21%)	37(82,22%)	64(63,37%)
Respuesta errónea	E_1	5(8,93%)	2(4,44%)	7(6,93%)
	E_2	1(1,79%)	0	1(0,99%)
	E_3	8(14,29%)	4(8,88%)	12(11,88%)
Total		56(100%)	45(100%)	101(100%)

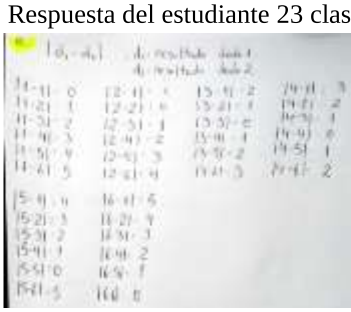
Nota. Elaboración propia.

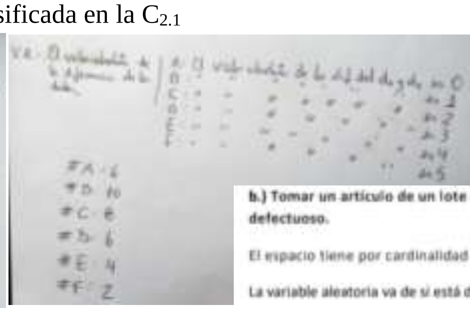
Según expone la Tabla 26, alrededor del 3% de la muestra respondió correctamente al $I_{1.3}$, dado que 0,99% de los participantes lograron identificar los valores de las dos variables aleatorias descritas previamente (a: 0, 1, 2, 3, 4 y 5; b: 0 y 1, donde 1 significa defectuoso y 0 no defectuoso) y representar cada una en un diagrama ($C_{2.2}$), y el 1,98% de los estudiantes establecieron el recorrido de las dos variables aleatorias definidas previamente y resentantaron graficamente sólo una de estas ($C_{2.1}$). La Figura 29 ejemplifica estos tipos de respuesta.

Figura 29

Ejemplos de respuestas correctas en $I_{1.3}$

Respuesta del estudiante 23 clasificada en la $C_{2.1}$





b.) Tomar un artículo de un lote de 20 artículos de una fábrica, para determinar si está o no defectuoso.

El espacio tiene por cardinalidad 20.

La variable aleatoria va de si está defectuoso o no, (defectuoso, no defectuoso)

Respuesta del estudiante 17 clasificada en la $C_{2.2}$

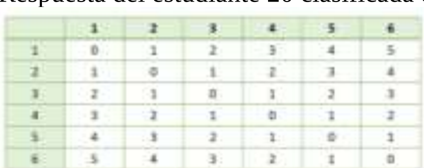
Nota. Elaboración propia.

Además en el $I_{1.3}$, el 13,86% de la muestra respondió parcialmente correcto, ya que el 7,92% de los participantes establecieron el recorrido de una o dos variables aleatorias anteriormente descritas ($C_{1.1}$) y el 5,94% de los estudiantes identificaron los valores de una de ellas y la representaron en lenguaje gráfico o tabular ($C_{1.2}$). Ejemplos de esta categoría de respuesta se muestran en la Figura 30.

Figura 30

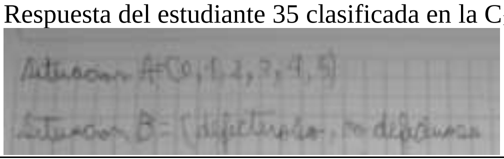
Ejemplos de respuestas parcialmente correctas en el $I_{1.3}$

Respuesta del estudiante 20 clasificada en la $C_{1.1}$



	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

Respuesta del estudiante 35 clasificada en la $C_{1.1}$



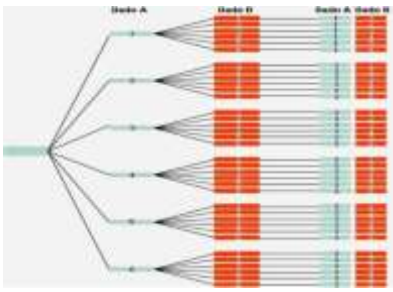
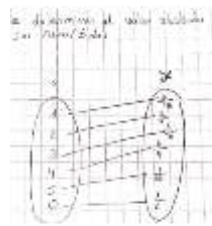
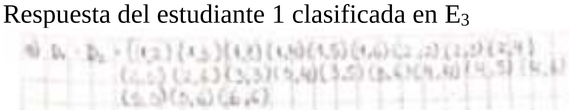
Respuesta del estudiante 85 clasificada en la $C_{1.2}$

Nota. Elaboración propia.

Mientras que sobre el 60% de la muestra no respondió al $I_{1.3}$ o escribió no saber lo que se pregunta (C_0). También alrededor del 20% de los participantes respondieron de forma errónea: el 6,93% de los estudiantes manifestaron confunsión entre el dominio de una variable aleatoria (el espacio muestral del experimento aleatorio involucrado) y su recorrido (conjunto de valores que puede tomar la variable aleatoria) (E_1); el 0,99% representaron la función de probabilidad en vez de su variable aleatoria (E_2); y el 11,88% de los jóvenes restantes propusieron algunos o todos los elementos del espacio muestral asociado al experimento aleatorio lanzar una moneda repetido cuatro veces (y/o vinculado al experimento aleatorio lanzar dos dados) (E_3). La Figura 31 ejemplifica los errores observados.

Figura 31

Ejemplos de respuestas errónea en el $I_{1.3}$

Respuesta del estudiante 5 clasificada en E_1 	Respuesta del estudiante 3 clasificada en E_2 
Respuesta del estudiante 86 clasificada en E_3 	Respuesta del estudiante 1 clasificada en E_3 

Nota. Elaboración propia.

En relación con los objetos matemáticos primarios involucrados en la solución del ítem 1, la mayoría de ellos han estado presentes en las respuestas correctas o parcialmente correctas de los participantes, pero en una baja proporción (ver Tabla 27). Alrededor del 43% de los estudiantes lograron diferenciar entre variables aleatorias y variables con dependencia funcional (S-$P_{1.1}$), aunque un bajo porcentaje definió aquellas variables aleatorias involucradas en las situaciones propuestas (S-$P_{1.2}$) y reconoció su correspondiente recorrido (S-$P_{1.4}$).

El lenguaje utilizado por los participantes en el ítem 1 principalmente ha sido verbal (60%), mediante expresiones verbales empleadas para seleccionar entre las situaciones dadas aquellas que pueden ser variables aleatorias y definir estas. Sólo el 16,83% de los jóvenes usaron símbolos para establecer el dominio y recorrido de estas variables aleatorias cuando: i) definieron un conjunto por extensión (ejemplo A= {0,1,2,3,4,5}) o por comprensión (B= {El valor absoluto de la diferencia entre d_1 y d_2 es 1}) y determinaron su cardinalidad (#B=10); ii) representaron el espacio muestral (Ω). También, pocos estudiantes emplearon el lenguaje gráfico (9%) para representar una variable aleatoria a través de un diagrama de árbol (ver Figura 29) o seudo diagrama sagital (ver Figura 30), y el lenguaje tabular (1%) para mostrar el recorrido de esta mediante una tabla de doble entrada (ver Figura 30),

Tabla 27

Frecuencia y porcentaje de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 1($I_{1.1}$ n=60 e $I_{1.3}$ n=17)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situaciones-Problemas	S-$P_{1.1}$ Diferenciar entre variables aleatorias y variables con dependencia funcional	43(42,57%)
	S-$P_{1.2}$ Definir variables aleatorias finitas en experimentos aleatorios	5(4,95%)
	S-$P_{1.4}$ Identificar recorrido o valores de una variable aleatoria finita	17(16,83%)
Lenguaje	Verbal	60(59,41%)
	Simbólico	17(16,83%)
	Gráfico	9(8,91%)
	Tabular	1(0,99%)
Conceptos	Experimento aleatorio	17(16,83%)
	Suceso aleatorio	17(16,83%)
	Espacio muestral	9(8,91%)
	Función (dominio y recorrido)	9(8,91%)
	Conjunto finito	17(16,83%)
	Variable aleatoria	43(42,57%)
	Variable dependiente e independiente	43(42,57%)
Proposiciones	Caracterización de la variable aleatoria y variables con dependencia funcional	43(42,57%)
	Caracterización de variables aleatorias discreta y continua	17(16,83%)
Procedimientos	Construcción del espacio muestral	9(8,91%)
	Partición disjunta del espacio muestral	9(8,91%)
	Construcción de un diagrama	9(8,91%)
Argumento	Verbal-deductivo	0
	Mediante representación gráfica	9(8,91%)

Nota. Elaboración propia.

El concepto de variable aleatoria ha sido utilizado por alrededor del 43% de los participantes, aunque los conceptos vinculados a esta, como función (8,91%), experimento aleatorio (16,83%) y espacio muestral (8,91%), fueron aplicados en menor proporción. Las dos proposiciones involucradas en las tareas han sido empleadas implícitamente por los estudiantes, estos mayormente utilizaron la relacionada con las características de una variable aleatoria y variables con dependencia funcional (42,57%) que la correspondiente a la diferencia entre variables aleatorias de tipo discretas y continuas (16,83%). Los procedimientos necesarios para resolver las situaciones-problemas propuestas fueron desarrollado por pocos participantes (8,91%): construyeron el espacio muestral, particionaron este según la variable aleatoria definida, y construyeron un diagrama para representarla.

Por último, aunque algunos participantes obtuvieron una respuesta correcta o parcialmente correcta y evidenciaron el dominio de ciertos objetos matemáticos primarios, ellos no lograron argumentar su respuesta justificando la aplicación de las dos proposiciones involucradas (razonamiento verbal- deductivo). Además, pocos estudiantes validaron su respuesta utilizando una representación gráfica como un diagrama (8,91%).

5.2.2.2 Resultados en el ítem 2.1

El ítem 2.1 presentado en la Figura 32 está compuesto por el $I_{2.1a}$ e $I_{2.1b}$ extraídos de Flores et al. (2014), que evalúan respectivamente la S-$P_{2.1}$ determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial, y la S-$P_{2.4}$ representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta (Bizet et al., 2023c, p.187)

Figura 32

Ítem 2.1 del cuestionario

2.1 Llamaremos lanzamiento a la acción de lanzar tres monedas al aire al mismo tiempo (de preferencia de la misma denominación). Ahora imagínate que se realizan 1000 lanzamientos y en cada uno de ellos se observa la variable "número de caras que ocurren".

a) En la siguiente tabla anota en la fila superior, los posibles valores de la variable y en la fila inferior, el número de veces (o frecuencia) que crees que ocurra cada valor:

Caras					Total
Frecuencia					

b) Anota la probabilidad que asignas a la ocurrencia de cada valor de la variable:

Caras	0	1	2	3	Total
Probabilidad					

Nota. Extraído de Flores et al. (2014).

En el $I_{2.1a,}$ el análisis de las respuestas de los participantes permitió obtener las categorías y frecuencias expuestas en la Tabla 28.

Tabla 28

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{2.1a}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	2(3,57%)	1(2,22%)	3(2,97)
	$C_{2.2}$	7(12,5%)	1(2,22%)	8(7,92%)
Respuesta parcialmente correcta	$C_{1.1}$	13(23,21%)	7(15,56%)	20(19,8%)
	$C_{1.2}$	6(10,71%)	0	6(5,94%)
Respuesta en blanco	No aplica	21(37,5%)	32(71,11%)	53(52,48%)
Respuesta errónea	E_1	6(10,71%)	2(4,44%)	8(7,92%)
	E_2	1(1,79%)	2(4,44%)	3(2,97%)
Total		56(100%)	45(100%)	101(100%)

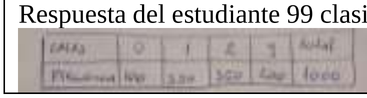
Nota. Elaboración propia.

Como exhibe la Tabla 28, alrededor del 11% de la muestra respondió correctamente al $I_{2.1a}$: el 7,92% de los participantes establecieron el recorrido de la variable (0,1,2,3), los valores teóricos de su distribución de frecuencias, en términos de valores absolutos, (125, 375, 375 y 125 respectivamente) y su suma corrresponde al tamaño de la muestra, igual a 1000 ($C_{2.2}$); y el 2,97% de los estudiantes determinaron los cuatro valores de la variable dada, la frecuencia (absoluta) de cada uno que varían levemente de su valor teórico y su suma ($C_{2.1}$). La Figura 33 muestra ejemplos de esta categoría de respuesta.

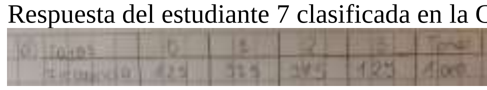
Figura 33

Ejemplos de respuestas correctas en el $I_{2.1a}$

Respuesta del estudiante 99 clasificada en la $C_{2.1}$



Respuesta del estudiante 7 clasificada en la $C_{2.2}$



Nota. Elaboración propia.

Además, el 26% de la muestra respondió parcialmente correcto, ya que el 19,8% de los participantes identificaron los cuatro valores de la variable ($C_{1.1}$) y el 5,94% de los estudiantes, además de lo descrito anteriormente, establecieron la suma de las frecuencias (absolutas) correspondiente a 1000 ($C_{1.2}$). Sin embargo, sólo algunos jóvenes, cuyas respuestas han sido clasificadas en esta categoría, lograron reconocer que los valores centrales de la variable (1 y 2) poseen mayor frecuencia. Ejemplos de la presente categoría de respuesta se exponen en la Figura 34.

Figura 34

Ejemplos de respuestas parcialmente correctas en el $I_{2.1a}$

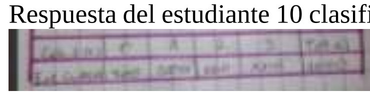
Respuesta del estudiante 32 clasificada en la $C_{1.1}$

Caras	0	1	2	3	Total
Frecuencia	240	153	302	53	748

Respuesta del estudiante 14 clasificada en la $C_{1.1}$

Cara	0	1	2	3	total
Frecuencia	100	300	300	100	800

Respuesta del estudiante 10 clasificada en la $C_{1.2}$



Respuesta del estudiante 20 clasificada en la $C_{1.2}$

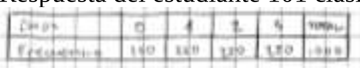
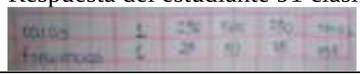
Caras	0	1	2	3	Total
Frecuencia	300	270	230	200	1000

Nota. Elaboración propia.

Mientras que sobre el 50% de la muestra no respondió o manifestó no saber lo que se pregunta en el $I_{2.1a}$ (C_0). También alrededor del 11% de los estudiantes respondieron erróneamente al ítem nombrado, debido a que: el 7,92% de los participantes afirmaron que los cuatro valores de la variable poseen igual frecuencia (E_1), y el 2,97% de los estudiantes registraron en la fila “caras” y “frecuencia” valores arbitrarios que no corresponden a los solicitados (E_2). La Figura 35 presenta ejemplos de los errores observados.

Figura 35

Ejemplo de respuesta errónea en el $I_{2.1a}$

Respuesta del estudiante 101 clasificada en E_1	Respuesta del estudiante 91 clasificada en E_2
	

Nota. Elaboración propia.

En cuanto al $I_{2.1b}$, la indagación de las respuestas proporcionada por los estudiantes permitió determinar las categorías y frecuencias presentadas en la Tabla 29.

Tabla 29

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{2.1b}$

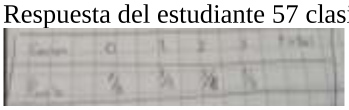
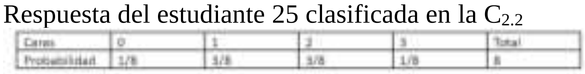
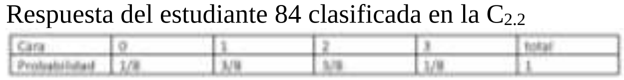
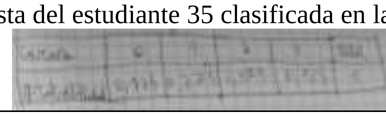
Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	3(5,36%)	1(2,22%)	4(3,96%)
	$C_{2.2}$	10(17,86%)	0	10(9,9%)
Respuesta parcialmente	$C_{1.1}$	7(12,5%)	0	7(6,93%)
correcta	$C_{1.2}$	2(3,57%)	0	2(1,98%)
Respuesta en blanco	No aplica	19(33,93%)	36(80%)	55(54,46%)
Respuesta errónea	E_1	6(10,71%)	1(2,22%)	7(6,93%)
	E_2	9(16,07%)	7(15,56%)	16(15,84%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Según presenta la Tabla 29, el 14% de la muestra respondió correctamente en el $I_{2.1b}$, debido a que el 3,96% de lo estudiantes establecieron para cada valor de la variable aleatoria discreta (0, 1, 2 y 3) su probabilidad asociada (1/8 o 12,5%; 3/8 o 37,5%; 3/8; y 1/8 respectivamente) ($C_{2.1}$), y el otro 9,9% de los participantes, además de lo señalado anteriormente, determinaron la probabilidad total correspondiente a 1 o 100% ($C_{2.2}$). Ejemplos de la presenta categoría de respuesta se exponen en la Figura 36.

Figura 36

Ejemplos de respuestas correctas en el $I_{2.1b}$

Respuesta del estudiante 57 clasificada en la $C_{2.1}$  Respuesta del estudiante 25 clasificada en la $C_{2.2}$ 	Respuesta del estudiante 84 clasificada en la $C_{2.2}$  Respuesta del estudiante 35 clasificada en la $C_{2.2}$ 

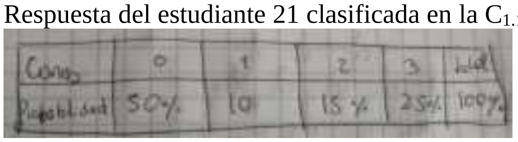
Nota. Elaboración propia.

Tambien alrededor del 9% de la muestra respondió parcialmente correcto al $I_{2.1b}$. El 6,93% de los participantes identificaron que la suma de las probabilidades es 1 o 100% ($C_{1.1}$), y el 1,98% de los estudiantes restantes, además de los expuesto anteriormente, calcularon adecuadamente la probabilidad vinculada a dos valores de la variable aleatoria propuesta y el total ($C_{1.2}$). La Figura 37 ejemplifica estos tipos de respuesta.

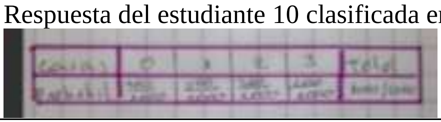
Figura 37

Ejemplos de respuestas parcialmente correctas en el $I_{2.1b}$

Respuesta del estudiante 21 clasificada en la $C_{1.1}$

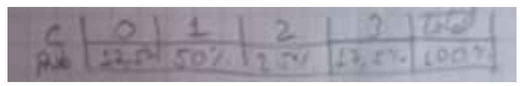


Respuesta del estudiante 10 clasificada en la $C_{1.1}$



Respuesta del estudiante 97 clasificada en la $C_{1.2}$

Respuesta del estudiante 12 clasificada en la $C_{1.2}$



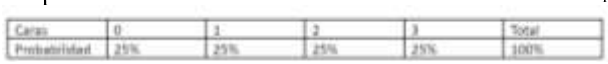
Nota. Elaboración propia.

Sin embargo más de la mitad de la muestra (54,46%) no respondió o expuso no saber lo que se pregunta en el $I_{2.1b}$ (C_0). Además alrededeador del 23% de los estudiantes respondieron de manera incorrecta al ítem señalado: el 6,93% de los participantes manifestaron en su razonamiento el sesgo de equiprobabilidad al afirmar que cada uno de los valores de la variable aleatoria discreta posee igual probabilidad (E_1) y; el 15,84% de los estudiantes propusieron valores arbitrarios que no corresponden a las probabilidades solicitadas y aquellos valores no suman 1 o 100% (E_2). Ejemplos de los errores mencionados se expone en la Figura 38.

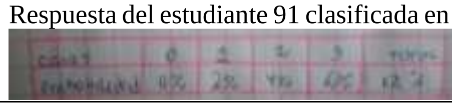
Figura 38

Ejemplos de respuesta errónea en el $I_{2.1b}$

Respuesta del estudiante 5 clasificada en E_1



Respuesta del estudiante 91 clasificada en E_2



Nota. Elaboración propia.

Sobre los objetos matemáticos primarios asociados a la solución del ítem 2.1, todos ellos se manifestaron en las respuestas correctas o parcialmente correctas de los participantes, aunque en un bajo porcentaje (ver Tabla 30). Solamente el 10,89% de los jóvenes hace una identificación (implícita) de probabilidades desde un enfoque frecuencial ($S\text{-}P_{2.1}$): determinaron para cada valor de la variable (0,1,2 y 3) su frecuencia absoluta y el valor total de la muestra 1000. También, una baja proporción (15,84%) de los estudiantes logró representar en una tabla la función de probabilidad (S-$P_{2.4}$).

El lenguaje empleado por los participantes en el ítem 2.1 en su mayoría fue verbal (36,63%), utilizado para entender el enunciado del problema. Además, pocos estudiantes han usado el lenguaje numérico (10,89%) para calcular la frecuencia absoluta o cuantificar la posibilidad de ocurrencia de un determinado valor de la variable involucrada. Aún más, sólo el 16,83% de los jóvenes utilizaron la tabla dada para presentar la probabilidad de cada valor de la variable aleatoria, y el 1,98% de los participantes ha usado el lenguaje simbólico, mediante el símbolo de porcentaje (%), cuando representaron cuantitativamente la probabilidad.

Tabla 30

Frecuencia y porcentaje de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 2.1($I_{2.1a}$ n=37 e $I_{2.1b}$ n=23)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situaciones-Problemas	S-$P_{2.1}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial	11(10,89%)
	S-$P_{2.4}$ Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta	16(15,84%)
Lenguaje	Verbal	37(36,63%)
	Numérico	11(10,89%)
	Simbólico	2(1,98%)
	Tabular	17(16,83%)
Conceptos	Experimento aleatorio	16(15,84%)
	Suceso aleatorio	16(15,84%)
	Espacio muestral	16(15,84%)
	Variable aleatoria	37(36,63%)
	Conjunto finito	16(15,84%)
	Enfoque frecuencial de probabilidad	11(10,89%)
	Enfoque clásico de probabilidad	16(15,84%)
	Distribución de frecuencias	17 (16,83%)
	Función de probabilidad	16(15,84%)
Proposiciones	Caracterización de variables aleatorias discreta y continua	30(29,7%)
	Convergencia, al crecer el número de ensayos la frecuencia relativa se va estabilizando	11(10,89%)
	Caracterización de la variable aleatoria mediante su función de probabilidad	16(15,84%)
	Propiedades de la función de probabilidad	19(18,81%)
	Regla de Laplace	16(15,84%)
Procedimientos	Partición disjunta del espacio muestral	16(15,84%)
	Reproducción de experimento aleatorio conservando las condiciones iniciales	17(16,83%)
	Cálculo de frecuencias absolutas	11(10,89%)
	Cálculo de probabilidades con la regla de Laplace	16(15,84%)
Argumento	Verbal-deductivo	11(10,89%)

Nota. Elaboración propia.

El concepto de variable aleatoria fue aplicado por alrededor del 37% de los estudiantes. Sin embargo, algunos conceptos relacionados con ella, como el enfoque de probabilidad clásico (15,84%) y frecuencial (10,89%), su función de probabilidad (15,84%) y distribución de frecuencias (16,83%), han sido empleado en un menor porcentaje. Las cinco proposiciones asociadas a las tareas fueron utilizadas implícitamente por los participantes, quienes principalmente aplicaron la proposición relacionada con diferenciar entre variables aleatorias de tipo discretas y continuas (29,7%) y las propiedades de la función de probabilidad (18,81%), a diferencia de la proposición sobre convergencia (10,89%) y la regla de Laplace (15,84%).

En tanto que los procedimientos requeridos para solucionar las situaciones-problemas propuestas han sido desarrollados por pocos estudiantes, por ejemplo, el 10,89% de los jóvenes calcularon frecuencias absolutas y el 15,84% lograron calcular probabilidades asociadas a valores de una variable aleatoria discreta. Por último, solamente el 10,89% de los participantes validaron su respuesta de forma deductiva, a través del cálculo correcto de frecuencias y probabilidades.

5.2.2.3 Resultados en el ítem 2.2

El ítem 2.2 mostrado en la Figura 39 esta contituido por: el $I_{2.2a}$ extraído de Salazar (2014) que evalúa la S-$P_{2.2}$ determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico (Bizet et al., 2023c, p.187); el $I_{2.2b}$ tomado de Salazar (2014) que valora la S-$P_{2.3}$ definir la función de probabilidad de una variable aleatoria utilizando la terminología $P(X= x_i)$ (Bizet et al., 2023c, p.187); y el $I_{2.2c}$ elaborado por el autor que evalúa la S-$P_{2.5}$ representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta (Bizet et al., 2023c, p.187).

Figura 39

Ítem 2.2 del cuestionario

2.2 En un programa de televisión los concursantes hacen girar una flecha en una ruleta como en la figura y cada vez que la flecha se detiene en la zona de color rojo, se gana un premio y el concurso se trata de hacer girar la flecha 3 veces. En base a esta situación responde las siguientes preguntas:

a) ¿Cuál es la probabilidad de que un participante gane dos premios? Y ¿Cuál es la probabilidad de que un concursante no gane algún premio?

b) Escribe en la tabla siguiente la función de probabilidad de la variable aleatoria que relaciona el espacio muestral con el número de premios que gana un concursante.

X	0	1	2	3
f(x)				

c) En un gráfico represente la función de probabilidad anterior.

Nota. Extraído de Salazar (2014) y elaborado por el autor.

En el $I_{2.2a}$, el análisis de las respuestas entregadas por los participantes hizo posible establecer las categorías y frecuencias presentadas en la Tabla 31.

Tabla 31

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{2.2a}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	3(5,36%)	0	3(2,97%)
	$C_{2.2}$	2(3,57%)	0	2(1,98%)
Respuesta parcialmente	$C_{1.1}$	9(16,07%)	0	9(8,91%)
correcta	$C_{1.2}$	4(7,14%)	0	4(3,96%)
Respuesta en blanco	No aplica	7(12,5%)	5(11,11%)	12(11,88%)
Respuesta errónea	E_1	15(26,79%)	21(46,67%)	36(35,64%)
	E_2	16(28,57%)	19(42,22%)	35(34,65%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como se aprecia en la Tabla 31, el 5% de la muestra respondió correctamente en el $I_{2.2a}$ sin justificar su respuesta, debido a que el 2,97% de los participantes determinaron la probabilidad de ganar dos premios al girar la flecha de la ruleta tres veces, correspondiente a 0,1406 o 14,06%, y la probabilidad de ganar cero premio al girar la fecha de la ruleta tres veces igual a 0,4218 o 42,18% ($C_{2.2}$). El 1,98% de los estudiantes restantes establecieron el valor de una de aquellas probabilidades y propusieron un valor aproximado para la otra ($C_{2.1}$). La Figura 40 ejemplifica la presente categoría de respuesta.

Figura 40

Ejemplos de respuestas correctas en el $I_{2.2a}$

Respuesta del estudiante 20 clasificada en la $C_{2.1}$

la probabilidad de que un participante gane 2 premios es de 16% y la de que no gane ningún premio es del 42%

Respuesta del estudiante 23 clasificada en la $C_{2.1}$

La probabilidad de que gane 2 premios es de 0,140625 y que gane 0 premios es de 0,4375

Respuesta del estudiante 14 clasificada en la $C_{2.2}$

que gane dos premios es de 9/64 y no gane ningún premio 27/64

Nota. Elaboración propia.

Además alrededor del 13% de la muestra respondió parcialmente correcto al $I_{2.2a}$; el 8,91% de los participantes sólo establecieron adecuadamente la probabilidad de ganar cero premios al girar la flecha de la ruleta tres veces, pero sin argumentar su respuesta ($C_{1.1}$); y el 3,96% de los estudiantes calcularon lo señalado anteriormente y validaron su respuesta mediante la regla del producto ($C_{1.2}$). Ejemplos de esta categoría de respuesta se exponen en la Figura 41.

Figura 41

Ejemplos de respuestas parcialmente correctas en el $I_{2.2a}$

Respuesta del estudiante 19 clasificada en la $C_{1.1}$ que gane dos premios: 3/64 que no gane : 27/64 Respuesta del estudiante 86 clasificada en $C_{1.1}$ 1) La probabilidad que gane 2 premios es de 1/64 2) la probabilidad que no gane ningún premio es de 27/64
Respuesta del estudiante 17 clasificada en la $C_{1.2}$ Probabilidad de ganar dos premios: 1/4*1/4*3/4=4,6% Probabilidad de que no gane ningún premio: 3/4*3/4*3/4=42,1% Respuesta del estudiante 28 clasificada en la $C_{1.2}$ i)Dos premios: (1/4)*(1/4)=(1/16) ii) Ningun premio: (3/4)*(3/4)*(3/4)=(27/64)

Nota. Elaboración propia.

Aunque el 11,88% de la muestra no respondió o manifestó no saber lo que se pregunta en el $I_{2.2a}$ (C_0). Aún más sobre el 70% de los participantes respondieron erróneamente a dicho ítem, pues el 35,64% de los estudiantes consideraron girar una vez la flecha de la ruleta y calcularon la probabilidad de ganar premio y/o no ganar premio (E_1), y el 34,65% restante señalaron valores arbitrarios que no corresponden a las probabilidades solicitadas (E_2). La Figura 42 ejemplifica los errores observados.

Figura 42

Ejemplos de respuesta errónea en el $I_{2.2a}$

Respuesta del estudiante 77 clasificada en E_1 tiene una probabilidad de un 25% y tiene un 75% de probabilidad de que no gane ningún premio Respuesta del estudiante 78 clasificada en E_1 la probabilidad de que gane 2 premios es de 1/4 y la de que no gane ningún premio es de 3/4	
Respuesta del estudiante 76 clasificada en E_1 un 25 %? Respuesta del estudiante 51 clasificada en E_1 3/4	Respuesta del estudiante 32 clasificada en E_2 2/3, 0/3

Nota. Elaboración propia.

Por otro lado, la indagación de las respuestas al $I_{2.2b}$ ha permitido obtener las categorías y frecuencias expuestas en la Tabla 32.

Tabla 32

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{2.2b}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	1(1,79%)	0	1(0,99%)
	$C_{2.2}$	3(5,36%)	0	3(2,97%)
Respuesta parcialmente correcta	$C_{1.1}$	2(3,57%)	0	2(1,98%)
	$C_{1.2}$	7(12,5%)	0	7(6,93%)
Respuesta en blanco	No aplica	28(50%)	39(86,67%)	67(66,34%)
Respuesta errónea	E_1	4(7,14%)	1(2,22%)	5(4,95%)
	E_2	3(5,36%)	2(4,44%)	5(4,95%)
	E_3	8(14,29%)	3(6,66%)	11(10,89%)
Total		56(100%)	45(100%)	101(100%)

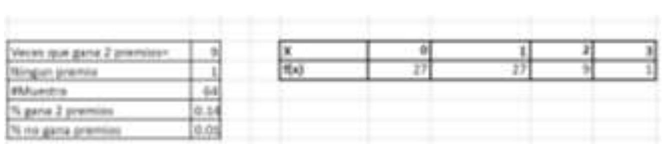
Nota. Elaboración propia.

Según exhibe la Tabla 32, el 4% de la muestra respondió correctamente al $I_{2.2b}$ sin argumentar su respuesta , ya que el 2,97% de los participantes determinaron para cada valor de la variable aleatoria X: número de premios que gana un concursante (0, 1, 2 y 3) su probabilidad sociada, correspondiente a ($C_{2.2}$): $\mathbf{f(0)} = \frac{\mathbf{27}}{\mathbf{64}} = \mathbf{0,4218\ o\ 42,18}\%$, $\mathbf{f(1)} = \frac{\mathbf{27}}{\mathbf{64}}, \mathbf{f(2)} = \frac{\mathbf{9}}{\mathbf{64}} = \mathbf{0,1406\ o\ 14,06}\%$ y $\mathbf{f(3)} = \frac{\mathbf{1}}{\mathbf{64}} = \mathbf{0,0156\ \ 1,56}$. Además el 0,99% de los estudiantes calcularon la cardinalidad del espacio muestral (64) asociado al experimento aleatorio girar la flecha de una ruleta repetido tres veces, y estableció el número de casos favorables del suceso aleatorio ganar cero premios (27), ganar un premio (27), ganar dos premios (9) y ganar tres premios (1) ($C_{2.1}$). Ejemplos de estos tipos de respuesta se exponen en la Figura 43.

Figura 43

Ejemplos de respuestas correctas en el $I_{2.2b}$

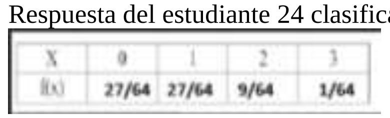
Respuesta del estudiante 23 clasificada en la $C_{2.1}$



Veces que gana 2 premios=	9
Ningun premio	1
#Muestra	64
% gana 2 premios	0.14
% no gana premios	0.05

X	0	1	2	3
f(x)	27	27	9	1

Respuesta del estudiante 24 clasificada en la $C_{2.2}$



X	0	1	2	3
f(x)	27/64	27/64	9/64	1/64

Respuesta del estudiante 26 clasificada en la $C_{2.2}$

X	0	1	2	3
f(x)	0,421875	0,421875	0,140625	0,015625

Nota. Elaboración propia.

También alrededor del 9% de la muestra respondió parcialmente correcto al $I_{2.2b}$ sin justificar su respuesta: el 1,98% de los participantes calcularon la probabilidad vinculada a un valor de la variable aleatoria involucrada, especificamente la probabilidad de que un participante gane tres premios al girar la flecha de una ruleta tres veces ($C_{1.1}$) ; y el 6,93% de los estudiantes determinaron la probabilidad asociado a dos valores de dicha variable, puntualmente la probabilidad de que un participante gane cero premios al girar la flecha de una ruleta tres veces, además de la probabilidad de que gane tres premio (o gane dos premios) ($C_{1.2}$). La Figura 44 ejemplifica esta categoría de respuesta.

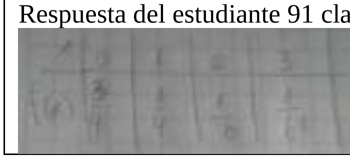
Figura 44

Ejemplos de respuestas parcialmente correctas en el $I_{2.2b}$

Respuesta del estudiante 5 clasificada en la $C_{1.1}$

x	0	1	2	3
f(x)	40%	25%	6.25%	1.56%

Respuesta del estudiante 91 clasificada en la $C_{1.1}$



Respuesta del estudiante 14 clasificada en la $C_{1.2}$

x	0	1	2	3
Probabilidad	27/64	9/64	9/64	27/64

Respuesta del estudiante 28 clasificada en la $C_{1.2}$

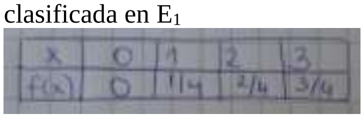
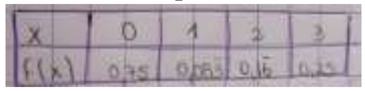
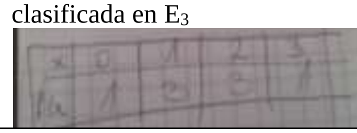
X	0	1	2	3
f(x)	27/64	16/64	20/64	1/64

Nota. Elaboración propia.

Por último, en el $I_{2.2b}$ más de la mitad de la muestra (66,34%) no respondió o señaló no saber lo que se pregunta (C_0). Aún más alrededor del 21% de la muestra respondió incorrectamente a aquel ítem: el 4,95% de los estudiantes asumieron una relación de linealidad entre la probabilidad del suceso (P) y los parámetros n (número de ensayos) y p (probabilidad de éxito) cuando emplearon incorrectamente la expresión $\mathbf{P = n \cdot p}$ para determinar la probabilidad requerida (E_1), pues su razonamiento ha sido, si la probabilidad de ganar premio al girar una vez la flecha de la ruleta es igual a $\mathbf{\frac{1}{4}}$, entonces la probabilidad de ganar premio al girar dos veces la flecha de la ruleta corresponde a $\mathbf{\frac{1}{4} \cdot 2}$ (de manera análoga para el otro caso $\mathbf{\frac{1}{4} \cdot 3}$); el 4,95% de los participantes calcularon la probabilidad de no ganar premio al girar una vez la flecha de la ruleta $\mathbf{\left(\frac{3}{4}\right)}$ en vez de girar tres veces $\mathbf{\left(\frac{27}{64}\right)}$ (E_2); y el 10,89% de los jóvenes restantes registraron en la tabla valores arbitrarios (E_3). Ejemplos de los errores observados se exponen en la Figura 45.

Figura 45

Ejemplos de respuesta errónea en el $I_{2.2b}$

Respuesta del estudiante 4 clasificada en E_1	Respuesta del estudiante 33 clasificada en E_2	Respuesta del estudiante 22 clasificada en E_3
		

Nota. Elaboración propia.

En cuanto al $I_{2.2c}$, el análisis de las respuestas entregadas por los estudiantes hizo posible determinar las categorías y frecuencias presentadas en la Tabla 33.

Tabla 33

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{2.2c}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	No aplica	1(1,79%)	0	1(0,99%)
Respuesta parcialmente correcta	$C_{1.1}$	2(3,57%)	0	2(1,98%)
	$C_{1.2}$	1(1,79%)	0	1(0,99%)
Respuesta en blanco	No aplica	35(62,5%)	44(97,78%)	79(78,22%)
Respuesta errónea	E_1	5(8,93%)	0	5(4,95%)
	E_2	1(1,79%)	0	1(0,99%)
	E_3	11(19,64%)	1(2,22%)	12(11,88%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como expone la Tabla 33 alrededor del 1% de los participantes respondieron correctamente en el $I_{2.2c}$ (C_2), debido a que representaron en un gráfico de barras los cuatro valores de la variable aleatoria discreta (0, 1, 2 y 3) con su correspondiente probabilidad ($\mathbf{f(0) = 0,4218}$; $\mathbf{f(1) = 0,4218}$; $\mathbf{f(2) = 0,1406}$ y $\mathbf{f(3) = 0,0156}$), pero omitieron el título del eje horizontal y vertical del gráfico. Ejemplo de este tipo de respuesta se muestra en la Figura 46.

Figura 46

Ejemplo de respuesta correcta en el $I_{2.2c}$

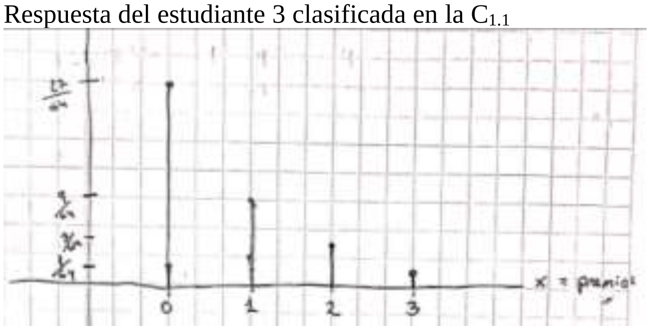
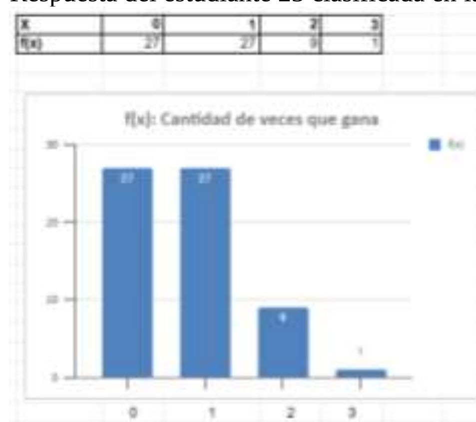
Respuesta del estudiante 26 clasificada en la $C_{2.2}$

Nota. Elaboración propia.

Además, el 3% de la muestra respondió parcialmente correcto al $I_{2.2c}$, ya que el 1,98% de los estudiantes representaron adecuadamente en un gráfico de barras sólo dos valores de la variable aleatoria (0 y 3) con su respectiva probabilidad $\left(\mathbf{f}(\mathbf{0})=\frac{\mathbf{27}}{\mathbf{64}}=\mathbf{0{,}4218}\ \mathbf{y}\ \mathbf{f}(\mathbf{3})=\frac{\mathbf{1}}{\mathbf{64}}=\mathbf{0{,}0156}\right)$, y señalaron el título del eje horizontal del gráfico ($C_{1.1}$). Mientras que el otro 0,99% de los participantes representaron gráficamente el recorrido de la variable aleatoria dada (0, 1, 2 y 3) y el número de sucesos aleatorios del espacio muestral asociado a cada valor (27, 27, 9 y 1 respectivamente), y asignaron título al eje vertical del gráfico. La Figura 47 ejemplifica la presente categoría de respuesta.

Figura 47

Ejemplos de respuestas parcialmente correctas en el $I_{2.2c}$

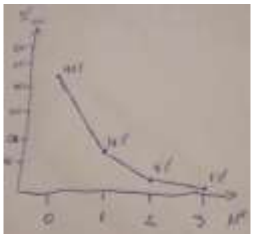
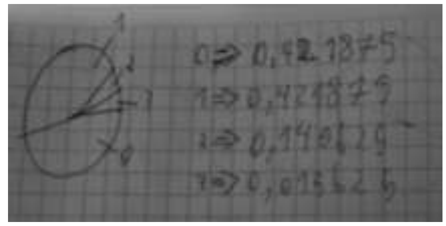
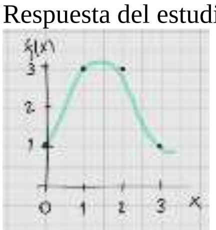
Respuesta del estudiante 3 clasificada en la $C_{1.1}$	Respuesta del estudiante 23 clasificada en la $C_{1.2}$
	

Nota. Elaboración propia.

Sin embargo gran parte de muestra (78,22%) no respondió o mencionó no saber lo que se pregunta en el $I_{2.2c}$ (C_0). Aún más, alrededor del 18% de los estudiantes respondieron de manera errónea a dicho ítem: el 4,95% de los participantes representaron en el plano cartesiano la función de probabilidad de la variable aleatoria discreta propuesta como si fuese una función continua en un intervalo de los números reales, aunque asignaron correctamente título a los ejes del gráfico (E_1); el 0,99% de los estudiantes representaron la función de probabilidad en un gráfico diferente al apropiado (gráfico de barras); y el 11,88% de los jóvenes restantes construyeron un gráfico sin sentido (E_3). Ejemplos de los errores observados se presentan en la Figura 48.

Figura 48

Ejemplos de respuesta errónea en el $I_{2.2c}$

Respuesta del estudiante 99 clasificada en E_1 	Respuesta del estudiante 35 clasificada en E_2 
Respuesta del estudiante 85 clasificada en E_3 	Respuesta del estudiante 98 clasificada en E_3 

Nota. Elaboración propia.

Respecto a los objetos matemáticos primarios involucrados en la solución del ítem 2.2, la mayoría estuvieron presentes en las respuestas correctas o parcialmente correctas de los estudiantes, pero en un bajo porcentaje (ver Tabla 34). Alrededor del 18% de los participantes cuantificaron la posibilidad de ocurrencia de algún valor de la variable aleatoria discreta propuesta desde el enfoque clásico de probabilidad (S-$P_{2.2}$). Mientras que una menor proporción de los jóvenes logró establecer su función de probabilidad en lenguaje simbólico (S-$P_{2.3}$) y representarla en lenguaje gráfico (S-$P_{2.5}$).

El lenguaje mayormente utilizado por los estudiantes ha sido tanto el numérico (17,82%), para cuantificar la posibilidad de ocurrencia, como el verbal (17,82%), para definir el suceso aleatorios compuestos correspondiente a cada valor de la variable aleatoria discreta, ejemplo: ganar 2 premios al girar la flecha de la ruleta 3 veces. Pocos participantes usaron: i) el lenguaje simbólico (11,88%), es decir, emplearon símbolos como f(x) y % cuando expresaron la probabilidad de ocurrencia de un suceso (compuesto); y ii) el lenguaje tabular (12,87%), o sea, utilizaron una tabla para presentar la probabilidad de cada valor de la variable aleatoria. Además, solamente el 3,96% de los jóvenes empleó el lenguaje gráfico para representar la función de probabilidad mediante un gráfico de barras.

Tabla 34

Frecuencia y porcentaje de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 2.2 ($I_{2.2a}$ n=18, $I_{2.2b}$ n=13 e $I_{2.2c}$ n=4)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situaciones-Problemas	S-$P_{2.2}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico	18(17,82%)
	S-$P_{2.3}$ Definir la función de probabilidad de una variable aleatoria utilizando la terminología P (X = x_i)	12(11,88%)
	S-$P_{2.5}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta	3(2,97%)
Lenguaje	Verbal	18(17,82%)
	Numérico	18(17,82%)
	Simbólico	12(11,88%)
	Gráfico	4(3,96%)
	Tabular	13(12,87%)
Conceptos	Experimento aleatorio	18(17,82%)
	Suceso aleatorio	18(17,82%)
	Sucesos independientes	18(17,82%)
	Espacio muestral	18(17,82%)
	Variable aleatoria	13(12,87%)
	Conjunto finito	18(17,82%)
	Enfoque clásico de probabilidad	18(17,82%)
	Función de probabilidad	13(12,87%)
	Combinatoria	5(4,95%)
	Distribución binomial	0
	Función de probabilidad de la binomial	0
Proposiciones	Caracterización de variables aleatorias discreta y continua	13(12,87%)
	Caracterización de la variable aleatoria mediante su función de probabilidad	3(2,97%)
	Propiedades de la función de probabilidad	3(2,97%)
	Caracterización de la distribución binomial	0
	La probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento	18(17,82%)
	Regla del producto	4(3,96%)
	Regla de Laplace	18(17,82%)
Procedimientos	Partición disjunta del espacio muestral	4(3,96%)
	Cálculo de la probabilidad (de éxito) mediante la regla de Laplace	18(17,82%)
	Cálculo de probabilidad compuesta con la regla del producto	4(3,96%)
	Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial	0
	Cálculo de combinaciones	5(4,95%)
	Construcción de un gráfico de barras	3(2,97%)
Argumento	Verbal-deductivo	4(3,96%)
	Mediante representación gráfica	3(2,97%)

Nota. Elaboración propia.

Los conceptos de variable aleatoria y función de probabilidad han sido aplicados por alrededor del 12,87% de los participantes, aunque algunos conceptos vinculados a ellas, como el enfoque de la probabilidad clásico (17,82%), suceso aleatorio (17,82%) y sucesos independientes (17,82%) han sido empleado en una mayor proporción. Las proposiciones relacionadas a las tareas fueron usadas implícitamente por los estudiantes. Ellos mayormente emplearon la proposición que indica que la probabilidad de éxito y la probabilidad de fracaso es constante (17,82%), aquella relacionada con diferenciar entre variables aleatorias de tipo discretas y continuas (12,87%) y la regla de Laplace (17,82%), en comparación con la proposición sobre caracterizar la variable aleatoria mediante su función de probabilidad (2,97%), la regla del producto (3,96%) y las propiedades de la función de probabilidad (2,97%).

En tanto que los procedimientos necesarios para solucionar las situaciones-problemas dadas fueron desarrollados por pocos jóvenes; i) el 3,96% de los estudiantes calcularon la probabilidad compuesta a través de la regla del producto; y ii) el 2,97% de los participantes construyeron un gráfico de barras para representar la función de probabilidad de una variable aleatoria discreta. Por tanto, sólo aquellos jóvenes validaron su respuesta mediante el razonamiento verbal-deductivo y mediante representación gráfica respectivamente.

5.2.2.4 Resultados en el ítem 3.1

El ítem 3.1 exhibido en la Figura 49 fue extraído de Muñoz et al. (2013) y valora la S-$P_{3.1}$ representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua (Bizet et al., 2023c, p.187).

Figura 49

Item 3.1 del cuestionario

3.1 Considera la función de densidad f(x) = 0,2 definida en el intervalo [0, 5] y construye la gráfica de f.

Nota. Extraído de Muñoz et al. (2013).

En el $I_{3.1}$, la indagación de las respuestas proporcionadas por los participantes ha permitido establecer las categorías y frecuencias expuestas en la Tabla 35.

Tabla 35

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{3.1}$

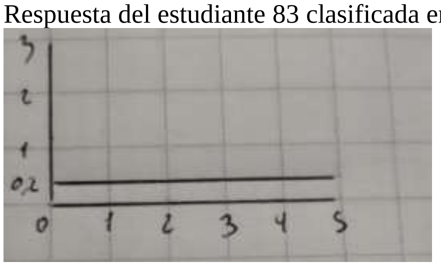
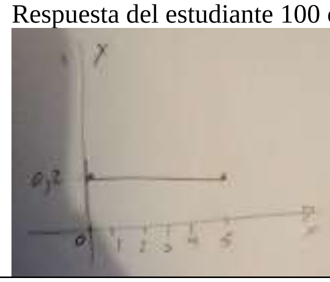
Categorías	Subcategoría	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	12(21,43%)	0	12(11,88%)
	$C_{2.2}$	2(3,57%)	0	2(1,98%)
Respuesta parcialmente	$C_{1.1}$	6(10,71%)	1(2,22%)	7(6,93%)
correcta	$C_{1.2}$	4(7,14%)	0	4(3,96%)
Respuesta en blanco	No aplica	29(51,79%)	40(88,89%)	69(68,32%)
Respuesta errónea	E_1	2(3,57%)	0	2(1,98%)
	E_2	1(1,79%)	4(8,89%)	5(4,95%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Según exhibe la Tabla 35, alrededor del 14% de la muestra respondió correctamente en el $I_{3.1}$, debido a que el 1,98% de los participantes representaron gráficamente la función de probabilidad de una variable aleatoria continua (o función de densidad) propuesta ($C_{2.2}$), y el 11,88% de los estudiantes graficaron en el plano cartesiano aquella función, pero no señalaron el título de los ejes X e Y del gráfico ($C_{2.1}$). Ejemplos de esta categoría de respuesta se muestran en la Figura 50.

Figura 50

Ejemplos de respuestas correctas en el $I_{3.1}$

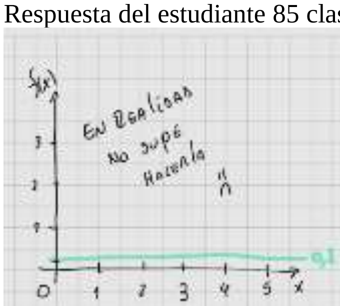
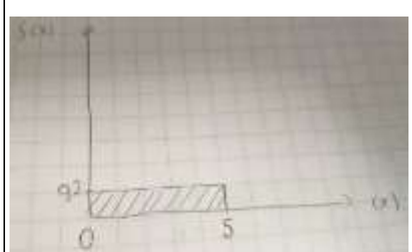
Respuesta del estudiante 83 clasificada en la $C_{2.1}$	Respuesta del estudiante 100 clasificada en la $C_{2.2}$
	

Nota. Elaboración propia.

También el 11% de la muestra respondió parcialmente correcto al $I_{3,1}$, donde el 3,96% de los estudiantes graficaron la función de densidad dada además de señalar la superficie correspondiente a $\mathbf{P(0 \leq X \leq 5) = 1}$ ($C_{1.2}$) y el 6,93% de los participantes representaron la función $\mathbf{f(x) = 0,2}$ pero en el intervalo $[\mathbf{0}, +\infty[$ ($C_{1.1}$). La Figura 51 ejemplifica estos tipos de respuesta.

Figura 51

Ejemplos de respuestas parcialmente correctas en el $I_{3.1}$

Respuesta del estudiante 85 clasificada en la $C_{1.1}$	Respuesta del estudiante 30 clasificada en la $C_{1.2}$
	

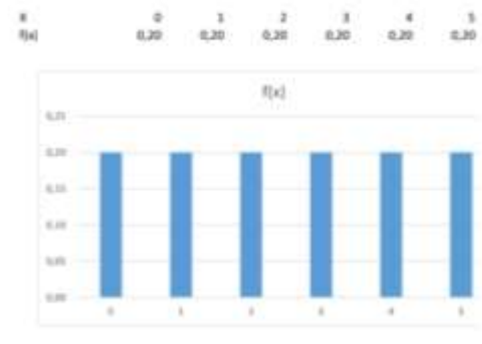
Nota. Elaboración propia.

Aunque gran parte de los participantes (69%) no respondieron o señalaron no saber lo que se pregunta en el $I_{3.1}$ (C_0). Además, alrededor del 7% de la muestra respondió incorrectamente a dicho ítem, debido a que el 1,98% de los estudiantes confundieron la función de densidad con la función de probabilidad de una variable aleatoria discreta y omitieron el título del eje horizontal y vertical del gráfico (E_1) y el 4,95% de los participantes construyeron una representación gráfica sin sentido (E_2). Ejemplos de los errores evidenciados se muestran en la Figura 52.

Figura 52

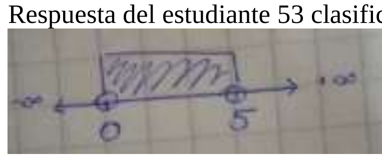
Ejemplos de respuestas erróneas en el $I_{3.1}$

Respuesta del estudiante 17 clasificada en E_1



Respuesta del estudiante 77 clasificada en E_2

Respuesta del estudiante 53 clasificada en E_2



Nota. Elaboración propia.

En relación con los objetos matemáticos primarios movilizados en la solución del ítem 3.1, todos ellos se han manifestado en respuestas correctas o parcialmente correctas de los estudiantes, aunque en una baja proporción (ver Tabla 36). Alrededor del 18% de los participantes lograron representar la función de probabilidad de una variable aleatoria continua en lenguaje gráfico (S-$P_{3.1}$). El lenguaje empleado por los jóvenes para entender el enunciado de la situación-problema propuesta ha sido tanto el verbal (24,75%), a través de términos (gráfica, función de densidad e intervalo), como el lenguaje simbólico (24,75%), por medio de los símbolos $\mathbf{f}(\mathbf{x})$ y $[\mathbf{0}, \mathbf{5}]$. Además el lenguaje gráfico fue utilizado por el 24,75% de los participantes cuando intentaron representar la función de densidad mediante una función constante en el plano cartesiano.

Tabla 36

Frecuencia y porcentaje de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 3.1(n=25)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situación-Problema	S-$P_{3.1}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua	18(17,82%)
Lenguaje	Verbal	25(24,75%)
	Simbólico	25(24,75%)
	Gráfico	25(24,75%)
Conceptos	Variable aleatoria	25(24,75%)
	Función (dominio y recorrido)	25(24,75%)
	Intervalo	18(17,82%)
	Función de densidad	18(17,82%)
Proposiciones	Caracterización de variables aleatorias discreta y continua	25(24,75%)

	Propiedades de la función de densidad	18(17,82%)
Procedimiento	Construcción de la gráfica de una función real en el plano cartesiano	14(13,86%)
Argumentos	Mediante representación gráfica	14(13,86%)

Nota. Elaboración propia.

Los conceptos de variable aleatoria y función real han sido empleados por alrededor del 25% de los estudiantes, pero conceptos relacionados con estas, como función de densidad (17,82%) e intervalo (17,82%) respectivamente, fueron usados en menor porcentaje. Las dos proposiciones involucradas en la tarea fueron aplicadas implícitamente por los participantes, ellos mayormente utilizaron la relacionada con caracterizar variables aleatorias de tipo discretas y continuas (24,75%), que las propiedades de la función de densidad (17,82%), específicamente el área limitada por su gráfica en su dominio y el eje de la abscisa es 1.

Sobre el procedimiento requerido para solucionar la tarea dada, el 13,86% de los jóvenes lograron construir en el plano cartesiano una función real ($f(x) = 0{,}2$) en el intervalo propuesto ($[0{,}5]$), así representaron gráficamente una función de densidad. Entonces solamente aquellos participantes justificaron su respuesta por medio de una representación gráfica.

5.2.2.5 Resultados en el ítem 4

El ítem 4 presentado en la Figura 53 está compuesto por: el $I_{4.a}$ elaborado por el autor y que evalúa la S-$P_{4.1}$ determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta y la S-$P_{4.2}$ definir la función de distribución de una variable aleatoria discreta (Bizet et al., 2023c, p.187); y el $I_{4.b}$ extraído de Chacón et al. (2018) que valora la S-$P_{4.3}$ representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta (Bizet et al., 2023c, p.187).

Figura 53

Ítem 4 del cuestionario

4. Para el gráfico que muestra la función de probabilidad de una variable aleatoria X:

a) Escribe su función de distribución
b) Construye un gráfico que represente su función de distribución.

Nota. Elaborado por el autor y extraído de Chacón et al. (2018).

En el $I_{4.a}$, el análisis de las respuestas entregadas por los estudiantes ha hecho posible determinar las categorías y frecuencias expuestas en la Tabla 37.

Tabla 37

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{4.a}$

Categorías	Frecuencia(Porcentaje)		Total
	Grupo1	Grupo 2	
Respuesta correcta	2(3,57%)	0	2(1,98%)
Respuesta parcialmente correcta	1(1,79%)	0	1(0,99%)
Respuesta en blanco	46(82,14%)	36(80%)	82(81,19%)
Respuesta errónea	7(12,5)	9(20%)	16(15,84%)
Total	56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como muestra la Tabla 37, alrededor del 2% de los estudiantes respondieron correctamente al $I_{4.a}$, pues calcularon la probabilidad acumulada hasta cada uno de los valores de la variable aleatoria discreta involucrada (2, 4, 6 y 8), es decir, $\mathbf{P(X \leq 2) = F(2) = 0,15 \text{ si } X \leq 2}$, $\mathbf{P(X \leq 4) = F(4) = 0,5 \text{ si } X \leq 4}$, $\mathbf{P(X \leq 6) = F(6) = 0,7 \text{ si } X \leq 6}$; y $\mathbf{P(X \leq 8) = F(8) = 1 \text{ si } X \leq 8}$ (C_2). La Figura 54 ejemplifica este tipo de respuesta.

Figura 54

Ejemplos de respuestas correctas en el $I_{4.a}$

Respuesta del estudiante 23 clasificada en la C_2

F(x)= 0,15 si x<= 2 ; 0,5 si x<=4 ; 0,7 si x<=6 ; 1 si x<=8

Respuesta del estudiante 26 clasificada en la C_2

f(x)= { 0,15 , x<=2 { 0,5 , x<=4 { 0,7 , x<=6 { 1 , x<=8

Nota. Elaboración propia.

También el 1% de los participantes respondieron parcialmente correcto en el $I_{4.a}$, debido a que determinaron la probabilidad acumulada hasta un valor de la variable aleatoria discreta propuesta (C_1). Ejemplo de la presente categoría de respuesta se expone en la Figura 55.

Figura 55

Ejemplos de respuestas parcialmente correctas en el $I_{4.a}$

Respuesta del estudiante 13 clasificada en la C_1

F(x) = P(X menor o igual x), por ejemplo f(x)=0,15 x⩽2

Nota. Elaboración propia.

Mientras que sobre el 80% de la muestra no respondió o señaló no saber lo que se pregunta en el $I_{4.a}$ (C_0), y 15,84% de los estudiantes respondieron de forma incorrecta a este ítem cuando señalaron valores y/o datos que no corresponden a los solicitados (E). La Figura 56 ejemplifica el error observado.

Figura 56

Ejemplos de respuestas erróneas en el $I_{4.a}$

Respuesta del estudiante 16 clasificada en E f(x)= P(X=x) = (0,15/4)+(0,35/4)+(0,20/4)+(0,30/4)=1
Respuesta del estudiante 100 clasificada en E f(x)=2x x siendo su respectiva frecuencia

Nota. Elaboración propia.

Por otra parte, la indagación de las respuestas al $I_{4.b}$, proporcionada por los participantes, permitió obtener las categorías y frecuencias exhibidas en la Tabla 38.

Tabla 38

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{4.b}$

Categorías	Subcategorías	Frecuencia(Porcentaje) Grupo1	Grupo 2	Total
Respuesta correcta	No aplica	1(1,79%)	0	1(0,99%)
Respuesta parcialmente correcta	No aplica	0	0	0
Respuesta en blanco	No aplica	49(87,5%)	43(95,56%)	92(91,09%)
Respuesta errónea	E_1	1(1,79%)	0	1(0,99%)
	E_2	1(1,79%)	0	1(0,99%)
	E_3	4(7,14%)	2(4,44%)	6(5,94%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Según expone la Tabla 38, el 1% de la muestra respondió correctamente al $I_{4.b}$, ya que representó graficamente la probabilidad acumulada hasta cada uno de los valores de la variable aleatoria discreta propuesta, es decir, graficó en el plano cartesiano una función escalonada, aunque omitió el título de los ejes del gráfico (C_2). La Figura 57 ejemplifica el presente tipo de respuesta. Además ningún participante respondió parcialmente correcto a dicho ítem, pues nadie representó graficamente la probabilidad acumulada hasta dos de los cuatro valores de la variable involucrada (C_1).

Figura 57

Ejemplo de respuesta correcta en el $I_{4.b}$

Respuesta del estudiante 36 clasificada en la C_2

Nota. Elaboración propia.

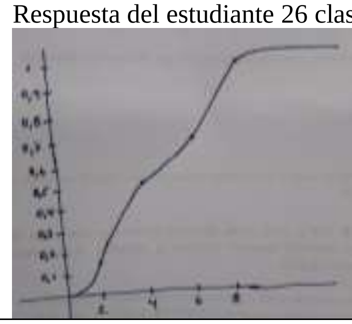
Por el contrario, la mayoría de los estudiantes (91,09%) no respondieron o mencionaron no saber lo que se pregunta en el $I_{4.b}$ (C_0). Aún más alrededor del 8% de la muestra respondió erróneamente a aquel ítem: el 0,99% de los participantes representaron en un gráfico de barras la probabilidad acumulada hasta cada uno de los valores de la variable aleatoria discreta involucrada, y no señalaron el título de los ejes del gráfico (E_1); el otro 0,99% graficaron la función de distribución de una variable aleatoria discreta como una función continua en un intervalo de los números reales, y omitieron el títulos de los ejes del gráfico (E_2); y el 5,94% restante construyeron un gráfico sin sentido (E_3). Ejemplos de los errores observados se exponen en la Figura 58.

Figura 58

Ejemplos de respuestas erróneas en el $I_{4.b}$

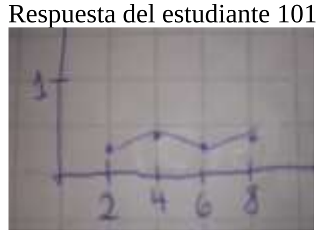
Respuesta del estudiante 23 clasificada en E_1

Respuesta del estudiante 26 clasificada en E_2



Respuesta del estudiante 62 clasificada en E_3

Respuesta del estudiante 101 clasificada en E_3



Nota. Elaboración propia.

Sobre los objetos matemáticos primarios vinculados a la solución del ítem 4, todos aquellos se han presentado en respuestas correctas o parcialmente correctas de los participantes, pero en un bajo porcentaje (ver Tabla 39). Alrededor del 3% de los estudiantes calcularon la probabilidad acumulada hasta alguno(s) de los valores de una variable aleatoria discreta (S-$P_{4.1}$). Mientras que una menor proporción de los jóvenes logró establecer su función de distribución en lenguaje simbólico (S-$P_{4.2}$) y representarla en lenguaje gráfico (S-$P_{4.3}$).

El lenguaje utilizado por los participantes para entender el enunciado de la tarea propuesta fue tanto el verbal (3,96%), por medio de términos (gráfico, función de probabilidad y variable aleatoria), como el gráfico (3,96%), a través de un gráfico de barras presentado. También el lenguaje gráfico ha sido empleado por un joven cuando representó una función de distribución mediante una función por parte en el plano cartesiano. Además, pocos estudiantes usaron: i) el lenguaje simbólico (2,97%), mediante símbolos como $F(x)$, $f(X \leq 2)$ y $X \leq 2$, cuando expresaron la probabilidad de ocurrencia de un suceso; y ii) lenguaje numérico (2,97%) para cuantificar la posibilidad de ocurrencia.

Tabla 39

Frecuencia y porcentaje de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 4 ($I_{4.a}$ n=3 e $I_{4.b}$ n=1)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situaciones-Problemas	S-$P_{4.1}$ Determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta	3(2,97%)
	S-$P_{4.2}$ Definir la función de distribución de una variable aleatoria discreta	2(1,98%)
	S-$P_{4.3}$ Representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta	1(0,99%)
Lenguaje	Verbal	4(3,96%)
	Numérico	3(2,97%)
	Simbólico	3(2,97%)
	Gráfico	4(3,96%)
Conceptos	Variable aleatoria	4(3,96%)
	Función (dominio y recorrido)	4(3,96%)
	Función de probabilidad	4(3,96%)
	Función de distribución	4(3,96%)
Proposiciones	Caracterización de variables aleatorias discreta y continua	4(3,96%)
	Caracterización de la variable aleatoria mediante su función de probabilidad	4(3,96%)
	Propiedad de la función de probabilidad	3(2,97%)
	Caracterización de la variable aleatoria mediante su función de distribución	4(3,96%)
Procedimientos	Construcción de la gráfica de una función real en el plano cartesiano	1(0,99%)
	Lectura e interpretación de un gráfico de barras	4(3,96%)
	Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de distribución	4(3,96%)
Argumentos	Mediante representación gráfica	4(3,96%)

Nota. Elaboración propia.

Los conceptos de variable aleatoria y función de distribución han sido aplicados por alrededor del 4% de los estudiantes, al igual que ciertos conceptos vinculados a ellas como función de probabilidad (3,96%) y función real (3,96%). Las cuatro proposiciones asociadas a las situaciones-problemas fueron utilizadas implícitamente por los participantes. Estos mayormente emplearon la proposición sobre diferenciar entre variables aleatorias de tipo discretas y continuas (3,96%) y aquellas relacionadas con caracterizar la variable aleatoria a través de su función de probabilidad (3,96%) y mediante su función de distribución (3,96%), a diferencia de la propiedad de la función de probabilidad (2,97%), específicamente la suma de las probabilidades asignadas a todos los valores de la variable aleatoria es igual a 1.

En tanto que los procedimientos necesarios para solucionar las tareas propuestas han sido desarrollados por pocos estudiantes; i) el 3,96% de los jóvenes primero leyó e interpreto el gráfico de barras dado y luego calculó probabilidades acumuladas usando la función de distribución; y ii) el 0,99% de los participantes construyeron una función escalonada en el plano cartesiano para representar gráficamente la función de distribución de una variable aleatoria discreta. Por tanto, sólo el 3,96% de los jóvenes justificaron su respuesta mediante representación gráfica, pues utilizaron las características visuales del gráfico de barras propuesto para justificar su solución o construyeron un gráfico.

5.2.2.6 Resultados en el ítem 5

El ítem 5 expuesto en la Figura 59 está integrado por: el $I_{5.2b}$ adaptado de Osorio et al. (2019) que valora la S-P $_{5.3}$ calcular la varianza de una variable aleatoria discreta (Bizet et al., 2023c, p.187); y el $I_{5.2c}$ elaborado por el autor que evalúa la S-P $_{5.2}$ interpretar la media o valor esperado de una variable aleatoria discreta y la S-$P_{5.5}$ interpretar la desviación estándar de una variable aleatoria discreta (Bizet et al., 2023c, p.187).

Figura 59

Ítem 5 del cuestionario

5.2 Se define la variable aleatoria X: número de mascotas que tiene un estudiante. La función de probabilidad asociada es:

x_i	0	1	2	3
$P(x_i)$	0,325	0,375	0,225	0,075

b) ¿Cuál es la varianza de la variable?

c) En el curso 3 medio de un liceo, el número medio de mascotas que tiene un estudiante es 1 y la desviación estándar es 0,95 ¿Cómo interpretarías cada medida?

Nota. Adaptado de Osorio et al. (2019) y elaborado por el autor.

En el $I_{5.2b}$, el análisis de las respuestas entregadas por los estudiantes hizo posible establecer las categorías y frecuencias expuestas en la Tabla 40.

Tabla 40

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{5.2b}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	No aplica	4(7,14%)	0	4(3,96%)
Respuesta parcialmente correcta	No aplica	4(7,14%)	0	4(3,96%)
Respuesta en blanco	No aplica	35(62,5)	37(82,22%)	72(71,29%)
Errónea	E_1	1(1,79%)	1(2,22%)	2(1,98%)
	E_2	12(21,43%)	7(15,56%)	19(18,81%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como se aprecia en la Tabla 40, el 4% de la muestra respondió correctamente al $I_{5.2b}$, aunque sin justificar su respuesta. Estos participantes obtuvieron que $\mathbf{Var}(\mathbf{X}) = \mathbf{0,8475}$ (C_2), por lo que se cree que calcularon la varianza de una variable aleatoria discreta (X: número de mascotas que tiene un estudiante) mediante su expresión algebraica $\left(\mathbf{Var}(\mathbf{X}) = \mathbf{E}(\mathbf{X^2}) - \left(\mathbf{E}(\mathbf{X})\right)^{\mathbf{2}}\right)$. La Figura 60 ejemplifica esta categoría de respuesta.

Figura 60

Ejemplo de respuesta correcta en el $I_{5.2b}$

Respuesta del estudiante 11 clasificada en la C_2

0,8475

Nota. Elaboración propia.

Además el 4% de los estudiantes respondieron parcialmente correcto en el $I_{5.2b}$, pero sin argumentar su respuesta. Ellos señalaron que $\mathbf{Var}(\mathbf{X}) = \mathbf{0,9}$ (C_1), por lo que se supone que calcularon la varianza de la variable aleatoria propuesta utilizando su expresión algebraica, pero no elevaron al cuadrado su segundo término ($\mathbf{Var}(\mathbf{X}) = \mathbf{E}(\mathbf{X^2}) - \mathbf{E}(\mathbf{X})$). Ejemplo de este tipo de respuesta se muestra en la Figura 61.

Figura 61

Ejemplo de respuesta parcialmente correcta en el $I_{5.2b}$

Respuesta del estudiante 12 clasificada en la C_1

0.9

Nota. Elaboración propia.

Sin embargo la mayoría de los participantes (71,29%) no respondieron o mencionaron no saber lo que se pregunta en el $I_{5.2b}$ (C_0). También alrededor del 21% de la muestra respondió incorrectamente a dicho ítem: el 1,98% de los estudiantes confundieron la varianza de la variable aleatoria con su moda al afirmar que la primera es igual a 1(E_1); y el otro 18,81% propusieron un valor arbitrario (E_2). La Figura 62 ejemplifica los errores observados.

Figura 62

Ejemplo de respuesta errónea en el $I_{5.2b}$

Respuesta del estudiante 15 clasificada en E_1	Respuesta del estudiante 24 clasificada en E_2
b) ¿Cuál es la varianza de la variable? 1	0,0525 Respuesta del estudiante 98 clasificada en E_2 la varianza es 1,75

Nota. Elaboración propia.

En cuanto al $I_{5.2c}$, la indagación de las respuestas dadas por los estudiantes ha permitido determinar las categorías y frecuencias presentadas en la Tabla 41.

Tabla 41

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{5.2c}$

Categorías	Subcategorías	Frecuencia(Porcentaje) Grupo1	Grupo 2	Total
Correcta	No aplica	0	0	0
Parcialmente correcta	$C_{1.1}$	1(1,79%)	0	1(0,99%)
	$C_{1.2}$	1(1,79%)	0	1(0,99%)
En blanco	No aplica	38(67,86%)	34(75,56%)	72(71,29%)
Errónea	E_1	7(12,5%)	2(4,44%)	9(8,91%)
	E_2	9(16,07%)	9(20%)	18(17,82)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Según exhibe la Tabla 41, ningún participante respondió correctamente en el $I_{5.2b}$ (C_2), ya que nadie logró interpretar la media (1) y desviación estándar (0,95) de la variable aleatoria propuesta (número de mascotas de un estudiante del curso 3 medio), es decir, no señaló que los estudiantes del curso 3 medio, en promedio, tienen una mascota, y que el número de mascotas de aquel curso se desvía, en promedio, 0,95 unidades respecto a la media 1.

Aunque alrededor del 2% de la muestra respondió parcialmente correcto al $I_{5.2b}$, debido a que el 0,99% de los estudiantes interpretaron la media afirmando que esta indica que los estudiantes del curso 3 medio, en promedio, tienen una mascota ($C_{1.1}$). El otro 0,99% de los participantes, además de lo señalado anteriormente, mencionaron la idea que la desviación estándar indica la dispersión de los datos respecto a la media, aunque no interpretaron la desviación estándar del caso particular propuesto ($C_{1.2}$). La Figura 63 ejemplifica la presente categoría de respuesta.

Figura 63

Ejemplos de respuestas parcialmente correctas en el $I_{5.2c}$

Respuesta del estudiante 86 clasificada en la $C_{1.1}$ 1 es el promedio de números de mascotas que tiene una persona.
Respuesta del estudiante 16 clasificada en la $C_{1.2}$ Los datos muestran que el promedio de mascotas por estudiantes es uno y que como la desviación estándar es cercana a uno entonces, más dispersos están los datos, lo que quiere decir que están más alejados del promedio.

Nota. Elaboración propia.

Mientras que sobre el 70% de la muestra no respondió o señaló no saber lo que se pregunta en el $I_{5.2b}$ (C_0). Aún más alrededor del 27% de los participantes respondieron incorrectamente a este ítem: el 8,91% de los estudiantes confundieron la media de la variable aleatoria con su moda al afirmar que 1 es el valor de la variable con mayor frecuencia (E_1) y el 17,82% restantes señalaron ideas que no corresponden a las solicitadas (E_2). Ejemplo de los errores observados se muestran en la Figura 64.

Figura 64

Ejemplos de respuesta errónea en el $I_{5.2b}$

Respuesta del estudiante 14 clasificada en E_1 la mayoría del curso de 3 medio de un liceo tiene 1 mascota
Respuesta del estudiante 3 clasificada en E_2 que hay varios alumnos que no tienen mascotas o que tiene dos, ya que la desviacion estandar es de casi 1 Respuesta del estudiante 18 clasificada en E_2 que los datos estan bien distribuidos, una desviacion estándar de 0,95 indica que hay una cantidad similar de estudiantes con 0,1 y 2 mascotas Respuesta del estudiante 82 clasificada en E_2 se puede interpretar que si hay muchos estudiantes con mascotas

Nota. Elaboración propia.

Con relación a los objetos matemáticos primarios asociados a la solución del ítem 5, la mayoría se han manifestado en las respuestas correctas o parcialmente correctas de los estudiantes, aunque en una baja proporción (ver Tabla 42). Alrededor del 4% de los participantes calcularon la varianza de una variable aleatoria discretra (S-P $_{5.3}$). Sin embargo sólo el 1,98% de los jóvenes lograron interpretar su media (S-P $_{5.2}$) y ninguno interpretó su desviación estándar (S-P$_{5.5}$).

El lenguaje empleado por los jóvenes para entender el enunciado de las tareas dadas ha sido: i) el verbal (7,92%), por medio de términos como variable aleatoria, función de probabilidad, varianza, media y desviación estándar; ii) el simbólico (7,92%), a través de los símbolos X, $P(x_i)$ y x_i ; y iii) el tabular (7,92%), mediante la tabla presentada que representa una función de probabilidad. También el lenguaje verbal fue utilizado por dos estudiantes cuando intentaron interpretar el valor propuesto para la media o desviación estándar, empleando términos como promedio, datos y datos dispersos. Además pocos participantes han usado el lenguaje numérico (3,96%) para establecer el valor de la varianza de una variable aleatoria discreta.

Tabla 42

Frecuencia de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 5 ($I_{5.2b}$ n=8 e $I_{5.2c}$ n=2)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situaciones-Problemas	S-P $_{5.2}$ Interpretar la media o valor esperado de una variable aleatoria discreta	2(1,98%)
	S-P $_{5.3}$ Calcular la varianza de una variable aleatoria discreta	4(3,96%)
	S-P$_{5.5}$ Interpretar la desviación estándar de una variable aleatoria discreta	0
Lenguaje	Verbal	8(7,92%)
	Numérico	4(3,96%)
	Simbólico	8(7,92%)
	Tabular	8(7,92%)
Conceptos	Suceso aleatorio	8(7,92%)
	Variable aleatoria	8(7,92%)
	Función de probabilidad	8(7,92%)
	Media	2(1,98%)
	Varianza	8(7,92%)
	Desviación estándar	1(0,99%)
Proposiciones	Caracterización de variables aleatorias discreta y continua	8(7,92%)
	Caracterización de la variable aleatoria mediante su función de probabilidad	8(7,92%)
	La media es un representante de los datos a partir de los que ha sido calculada y no tiene por qué ser igual a uno de los valores de los datos	2(1,98%)
Procedimientos	Lectura de una tabla de probabilidad	8(7,92%)
	Cálculo de medida de dispersión asociada a una variable aleatoria discreta empleando su expresión algebraica	8(7,92%)
Argumentos	Verbal-deductivo	0

Nota. Elaboración propia.

Los conceptos de variable aleatoria y varianza fueron utilizados por alrededor del 8% de los estudiantes, aunque conceptos relacionados con la primera de estas, como media (1,98%) y desviación estándar (0,99%), han sido usados en menor porcentaje. Las tres proposiciones asociadas a las tareas han sido empleadas implícitamente por los participantes, quienes mayormente aplicaron la proposición sobre diferenciar entre variables aleatorias de tipo discretas y continuas (7,92%) y aquella relacionada con caracterizar la variable aleatoria mediante su función de probabilidad (7,92%), a diferencia de la propiedad en torno a la media (1,98%)

En tanto que los procedimientos necesarios para solucionar las situaciones-problemas dadas fueron empleados por pocos jóvenes. Ellos leyeron la tabla de probabilidad propuesta (7,92%) y luego reemplazaron los valores de la variable aleatoria con su respectiva probabilidad en la expresión algebraica de la varianza (7,92%), pero sin mostrar estos procedimientos.

Aunque pocos participantes alcanzaron una respuesta correcta o parcialmente correcta y evidenciaron el dominio de ciertos objetos matemáticos primarios, aquellos no lograron validar su respuesta argumentando la utilización de las proposiciones involucradas (razonamiento verbal-deductivo), ni tampoco consiguieron justificar su respuesta con base en los procedimientos asociados a la tarea.

5.2.2.7 Resultados en el ítem 6

El ítem 6 exhibido en la Figura 65, está compuesto por: el $I_{6.c}$ e $I_{6.d}$ extraídos de Alvarado y Retamal (2014) que valoran la S-$P_{6.1}$ identificar situaciones que pueden modelizarse a través de una distribución binomial, la S-$P_{6.2}$ determinar los parámetros asociados a una distribución binomial (n, p y q) y la S-$P_{6.4}$ determinar probabilidades asociadas a una distribución binomial de forma manual (Bizet et al., 2023c, p.187); y el $I_{6.e}$ elaborado por el autor que evalúa la S-$P_{6.1}$ identificar situaciones que pueden modelizarse a través de una distribución binomial, la S-$P_{6.2}$ determinar los parámetros asociados a una distribución binomial y la S-$P_{6.3}$ calcular la media de una distribución binomial (Bizet et al., 2023c, p.187).

Figura 65

Ítem 6 del cuestionario

6. Un profesor aplica a sus estudiantes un cuestionario sorpresa con cinco ítems de opción múltiple. Uno de los estudiantes no ha estudiado el material del cuestionario, y por tanto decide contestar los cinco ítems al azar, adivinando las respuestas sin leer las preguntas ni las opciones de respuestas.
Hoja de respuestas del cuestionario. Instrucciones: encierre en un círculo la mejor respuesta de cada ítem

1. A B C
2. A B C
3. A B C
4. A B C
5. A B C

antes de ver las respuestas correctas del cuestionario y encontrar cómo le fue a este estudiante, se considerarán algunos hechos que podrían suceder si un cuestionario se contesta de esta forma.
c) ¿cuál es la probabilidad de elegir las respuestas acertadas de los cinco ítems?
d) ¿cuál es la probabilidad de elegir las respuestas incorrectas de los cinco ítems?
e) Calcula la media de las respuestas correctas

Nota. Extraído de Alvarado y Retamal (2014) y elaborado por el autor.

En el $I_{6.c}$, el análisis de las respuestas de los participantes permitió obtener las categorías y frecuencias expuestas en la Tabla 43.

Tabla 43

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{6.c}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	12(21,43%)	1(2,22%)	13(12,87%)
	$C_{2.2}$	2(3,57%)	0	2(1,98%)
Respuesta parcialmente correcta	No aplica	3(5,35%)	7(15,56%)	10(9,9%)
Respuesta en blanco	No aplica	22(39,29%)	11(24,44%)	33(32,67%)
Respuesta errónea	E_1	6(10,71%)	10(22,22%)	16(15,84%)
	E_2	1(1,79%)	0	1(0,99%)
	E_3	10 (17,86%)	16(35,56%)	26(25,74%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como se aprecia en la Tabla 43, alrededor del 15% de la muestra respondió correctamente en el $I_{6.c}$, dado que el 12,87% de los estudiantes determinaron la probabilidad de elegir las respuestas acertadas de los cinco ítems en un cuestionario sin argumentar su respuesta ($C_{2.1}$), la cual corresponde a $\mathbf{P(5)} = \frac{\mathbf{1}}{\mathbf{243}} = \mathbf{0,004}$ o $\mathbf{0,4}$%. El 1,98% de los participantes restantes establecieron dicha probabilidad justificando su resultado ($C_{2.2}$), mediante la aplicación de las características de la distribución binomial y su función de probabilidad $\left(\mathbf{P(5)} = \mathbf{C_5^5} \cdot \left(\frac{\mathbf{1}}{\mathbf{3}}\right)^{\mathbf{5}} \cdot \left(\frac{\mathbf{2}}{\mathbf{3}}\right)^{\mathbf{0}} = \mathbf{1} \cdot \frac{\mathbf{1}}{\mathbf{243}} = \frac{\mathbf{1}}{\mathbf{243}}\right)$ o la utilización de la regla del producto y el concepto de sucesos independientes $\left(\mathbf{P(5)} = \mathbf{C_5^5} \cdot \frac{\mathbf{1}}{\mathbf{3}} \cdot \frac{\mathbf{1}}{\mathbf{3}} \cdot \frac{\mathbf{1}}{\mathbf{3}} \cdot \frac{\mathbf{1}}{\mathbf{3}} \cdot \frac{\mathbf{1}}{\mathbf{3}} = \mathbf{1} \cdot \frac{\mathbf{1}}{\mathbf{243}} = \frac{\mathbf{1}}{\mathbf{243}}\right)$, aunque en ninguna de las dos estrategias el estudiante explicitó el cálculo de combinaciones ($\mathbf{C_5^5} = \mathbf{1}$). La Figura 66 ejemplifica estos tipos de respuesta.

Figura 66

Ejemplos de respuestas correctas en el $I_{6.c}$

Respuesta del estudiante 57 clasificada en la $C_{2.1}$ 1/243 Respuesta del estudiante 26 clasificada en la $C_{2.1}$ 0,0041152 Respuesta del estudiante 95 clasificada en la $C_{2.1}$ 0,4%	Respuesta del estudiante 28 clasificada en la $C_{2.2}$ (1/3)^5 Respuesta del estudiante 10 clasificada en la $C_{2.2}$ 1/3 *1/3 *1/3 *1/3 *1/3 = 1/243

Nota. Elaboración propia.

También alrededor del 10% de los particiantes respondieron parcialmente correcto al $I_{6.c}$, debido a que sólo determinaron la probabilidad de elegir la respuesta correctamente de un ítem individual (C_1), sin explicitar que establecieron el valor del parámetro p de la binomial $\left(\mathbf{p} = \frac{1}{3}\right)$. Ejemplos de esta categoría de respuesta se exponen en la Figura 67.

Figura 67

Ejemplos de respuestas parcialmente correctas en el $I_{6.c}$

Respuesta del estudiante 40 clasificada en la C_1	Respuesta del estudiante 22 clasificada en la C_1
la probabilidad de acertar es de 1/3	de un 1 en 3, es decir, un 33.3%

Nota. Elaboración propia.

Sin embargo, el 32,67% de los estudiantes no respondieron o señalaron no saber sobre lo preguntado en el $I_{6.c}$ (C_0). Aún más, en dicho ítem el 42,57% de la muestra respondió incorrectamente: el 15,84% de los participantes manifestaron indicios del sesgo de equiprobabilidad (E_1), al afirmar que la probabilidad de elegir las respuestas acertadas de los cinco ítems es igual al 20% o $\frac{1}{5}$, es decir, suponen que los sucesos aleatorios elegir una, dos, tres, cuatro, o cinco respuestas correctas tienen igual probabilidad; el 0,99% de los estudiantes asumieron una relación lineal entre la probabilidad del suceso (P) y los parámetros p (probabilidad de éxito) y n (número de ensayos) cuando emplearon incorrectamente la expresión $P = \frac{p}{n}$ para determinar la probabilidad requerida (E_2), debido a que su razonamiento ha sido, si la probabilidad de elegir la respuesta correctamente de un ítem es igual $\frac{1}{3}$ entonces la probabilidad de elegir las respuestas acertadas de los cinco ítems corresponde a $\frac{\frac{1}{3}}{5}$; y el 25,74% de los participantes propusieron un valor arbitrario (E_3). La Figura 68 ejemplifica los errores observados.

Figura 68

Ejemplos de respuestas erróneas en el $I_{6.c}$

Respuesta del estudiante 8 clasificada en E_1	Respuesta del estudiante 43 clasificada en E_3
20%	10%
Respuesta del estudiante 98 clasificada en E_1	Respuesta del estudiante 64 clasificada en E_3
1/5	50%
Respuesta del estudiante 30 clasificada en E_2	
en cada ítem la probabilidad de acertar es de 1/3, entonces seria de 1/15	

Nota. Elaboración propia.

Por otro lado, la indagación de las respuestas al $I_{6.d}$ posibilitó determinar las categorías y frecuencias presentadas en la Tabla 44.

Tabla 44

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{6.d}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	9(16,07%)	1(2,22%)	10(9,9%)
	$C_{2.2}$	2(3,57%)	0	2(1,98%)
Respuesta parcialmente correcta	No aplica	5(8,93%)	6(13,33%)	11(10,89%)
Respuesta en blanco	No aplica	24(42,86%)	15(33,33%)	39(38,61%)
Respuesta errónea	E_1	6(10,71%)	1(2,22%)	7(6,93%)
	E_2	10(17,86%)	22(48,89%)	32(31,68%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Según expone la Tabla 44, alrededor del 12% de la muestra respondió correctamente al $I_{6.d}$, debido a que el 9,9% de los participantes establecieron la probabilidad de elegir las respuestas incorrectas de los cinco ítems en un cuestionario sin justificar su resultado ($C_{2.1}$), el cual corresponde a $\mathbf{P(0)} = \frac{\mathbf{32}}{\mathbf{243}} = \mathbf{0{,}131}$ o $\mathbf{13{,}1}\%$. El otro 1,98% de los estudiantes determinaron aquella probabilidad argumentando su respuesta ($C_{2.2}$), mediante la utilización de la regla del producto $\left(\mathbf{P(0)} = \mathbf{C_0^5} \cdot \frac{\mathbf{2}}{\mathbf{3}} \cdot \frac{\mathbf{2}}{\mathbf{3}} \cdot \frac{\mathbf{2}}{\mathbf{3}} \cdot \frac{\mathbf{2}}{\mathbf{3}} \cdot \frac{\mathbf{2}}{\mathbf{3}} = \mathbf{1} \cdot \frac{\mathbf{32}}{\mathbf{243}} = \frac{\mathbf{1}}{\mathbf{243}}\right)$ o la aplicación de las características de la distribución binomial y su función de probabilidad $\left(\mathbf{P(0)} = \mathbf{C_0^5} \cdot \left(\frac{\mathbf{1}}{\mathbf{3}}\right)^{\mathbf{0}} \cdot \left(\frac{\mathbf{2}}{\mathbf{3}}\right)^{\mathbf{5}} = \mathbf{1} \cdot \frac{\mathbf{32}}{\mathbf{243}} = \frac{\mathbf{32}}{\mathbf{243}}\right)$, pero en las dos estrategias el estudiante no explicitó el cálculo de combinaciones ($\mathbf{C_0^5 = 1}$). Ejemplos de estos tipos de respuesta se aprecian en la Figura 69.

Figura 69

Ejemplos de respuestas correctas en el $I_{6.d}$

Respuesta del estudiante 57 clasificada en la $C_{2.1}$

32/243

Respuesta del estudiante 26 clasificada en la $C_{2.1}$

0,1316872

Respuesta del estudiante 17 clasificada en la $C_{2.1}$

13,16%

Respuesta del estudiante 10 clasificada en la $C_{2.2}$

como la probabilidad de acertar a una pregunta es de 1/3, la probabilidad de no acertar a una respuesta sería 2/3; la probabilidad de que no acierte ninguna buena de los 5 ítems es 2/3 * 2/3 *2/3* 2/3* 2/3 = 32 / 243

Respuesta del estudiante 28 clasificada en la $C_{2.2}$

(2/3)^5

Nota. Elaboración propia.

Además el 10,89% de los particiantes respondieron parcialmente correcto en el $I_{6.d}$, ya que sólo determinaron la probabilidad de seleccionar una respuesta incorrecta en un ítem ($C_{1.2}$), sin explicitar que establecieron el valor del parámetro q de la binomial $\left(\mathbf{q} = \frac{2}{3}\right)$. La Figura 70 ejemplifica la presente categoría de respuesta.

Figura 70

Ejemplos de respuestas parcialmente correctas en el $I_{6.d}$

Respuesta del estudiante 40 clasificada en la C_1	Respuesta del estudiante 22 clasificada en la C_1
la probabilidad de elegir respuestas incorrectas es de 2/3	2 de 3, un 66.7%

Nota. Elaboración propia.

Aunque alrededor del 39% de la muestra no respondió o mencionó no saber lo que se pregunta en el $I_{6.c}$ (C_0). También, en aquel ítem el 38,61% de los participantes respondieron erróneamente: el 6,93% de estudiantes mostraron indicios del sesgo de equiprobabilidad (E_1), al señalar que la probabilidad de elegir las respuestas incorrectas de los cinco ítems es igual al 20% o $\frac{1}{5}$, o sea, suponen que los sucesos aleatorios elegir una, dos, tres, cuatro, o cinco respuestas incorrectas poseen misma probabilidad; y el 31,68% de la muestra señaló un valor arbitrario (E_2). Ejemplos de aquellos errores observados se exhiben en la Figura 71.

Figura 71

Ejemplos de respuestas erróneas en el $I_{6.d}$

Respuesta del estudiante 8 clasificada en E_1	Respuesta del estudiante 58 clasificada en E_2
20%	4/5
Respuesta del estudiante 98 clasificada en E_1	Respuesta del estudiante 81 clasificada en E_2
1/5	5/5

Nota. Elaboración propia.

En cuanto al $I_{6.e}$, el análisis de las respuestas proporcionada por los participantes ha hecho posible establecer las categorías y frecuencias exhibidas en la Tabla 45.

Tabla 45

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{6.e}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	No aplica	1(1,79%)	0	1(0,99%)
Respuesta parcialmente correcta	No aplica	0	0	0
Respuesta en blanco	No aplica	43(76,79%)	39(86,67%)	82(81,19%)
Respuesta errónea	E_1	1(1,79%)	1(2,22%)	2(1.98%)
	E_2	4(7,14%)	1(2,22)	5(4,95%)
	E_3	7(12,5)	4(8,89%)	11(10,89%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como muestra la Tabla 45, alrededor del 1% de los estudiantes respondieron correctamente en el $I_{6.e}$ (C_2) , pero sin justificar su resultado. Ellos obtuvieron que $\mathbf{\mu = 1,6}$, por lo que se cree que calcularon la media de una variable aleatoria con distribución binomial mediante su expresión algebraica $\mathbf{\mu = n \cdot p = 5 \cdot \frac{1}{3} = 1,6}$. La Figura 72 ejemplifica esta categoría de respuesta.

Figura 72

Ejemplo de respuesta correcta en el $I_{6.e}$

Respuesta del estudiante 4 clasificada en la C_2
1.666

Nota. Elaboración propia.

También ninguno de los participantes respondió parcialmente correcto al $I_{6.3}$ (C_1), ya que no propusieron un valor próximo a la media de la variable aleatoria (1,5 o 1,7).

Mientras que sobre el 80% de los estudiantes no respondieron o escribieron no saber lo que se pregunta en el $I_{6.e}$ (C_0). Aún más el 17,82% de la muestra respondió de manera incorrecta: el 1,98% de los participantes confundieron la media de la variable aleatoria con su mediana al señalar que la primera es igual a 2 (E_1); el 4,95% de los estudiantes afirmaron que la media de las respuestas correctas del cuestionario corresponde a la midad del número de ítems del cuestionario (E_2), o sea, 2,5; y el 10,89% de los jóvenes restantes señalaron un valor arbitrario (E_3). La Figura 73 ejemplifica los errores observados.

Figura 73

Ejemplos de respuestas erróneas en el $I_{6.e}$

Respuesta del estudiante 62 clasificada en la E_1	Respuesta del estudiante 76 clasificada en E_2	Respuesta del estudiante 68 clasificada en E_3
2	2.5 la mitad?	3

Nota. Elaboración propia.

Sobre los objetos matemáticos primarios asociados a la solución del ítem 6, todos ellos se han manifestado en las respuestas correctas o parcialmente correctas de los participantes, pero en una bajo porcentaje (ver Tabla 46). Alrededor de 1% de los estudiantes identificaron que la situación propuesta puede modelizarse mediante una distribución binomial (S-$P_{6.1}$). Sin embargo ciertos jóvenes lograron determinar los valor de n, p y q (S-$P_{6.2}$), pero sin señalar que son parámetros de la binomial: i) el 1,98% identificó el número de repeticiones independientes del experimento (n); ii) el 11,88% estableció el valor de la probabilidad de éxito en cada experimento (p); y iii) el 12,87% determinó el valor de la probabilidad de fracaso en cada uno (q). Además el 14,85% de los participantes determinaron probabilidades vinculadas a una distribución binomial de forma manual (S-P$_{6.4}$) y alrededor del 1% lograron establecer el valor de su media (S-P$_{6.3}$).

El lenguaje empleado por los estudiantes fue: i) el verbal (24,75%), aplicado tanto para entender el enunciado de las tareas como para resolver estas, mediante el uso de términos y expresiones verbales, como por ejemplo, probabilidad, elegir las respuestas acertadas de los cinco ítems (suceso aleatorio compuesto) y acertar a una pregunta individual (suceso aleatorio simple), etc. ; y el ii) el numérico (24,75%), utilizado para cuantificar la posibilidad de ocurrencia de un suceso o establecer el valor de la media de una variable aleatoria discreta. Además el lenguaje simbólico ha sido utilizado por pocos estudiantes (4,95%), es decir, estos usaron el símbolo de porcentaje (%) cuando representaron cuantitativamente la probabilidad de un suceso aletorio.

Tabla 46

Frecuencia de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 6 ($I_{6.c}$ n=25, $I_{6.d}$ n=23 e $I_{6.e}$ n=1)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situaciones-problemas	S-P$_{6.1}$ Identificar situaciones que pueden modelizarse a través de una distribución binomial	1(0,99%)
	S-P$_{6.2}$ Determinar los parámetros asociados a una distribución binomial como n (número de ensayos), p (probabilidad de éxito) y q (probabilidad de fracaso)	n: 2(1,98%) p: 12(11,88%) q: 13(12,87%)
	S-P$_{6.3}$ Calcular la media de una distribución binomial	1(0,99%)
	S-P$_{6.4}$ Determinar probabilidades asociadas a una distribución binomial de forma manual	15(14,85%)
Lenguaje	Verbal	25(24,75%)
	Numérico	25(24,75%)
	Simbólico	5(4,95%)
Conceptos	Experimento aleatorio	25(24,75%)
	Espacio muestral	15(14,85%)
	Suceso aleatorio	25(24,75%)
	Variable aleatoria	1(0,99%)
	Enfoque de probabilidad clásico	15(14,85%)
	Sucesos independientes	25(24,75%)
	Distribución binomial	1(0,99%)
	Función de probabilidad de la binomial	1(0,99%)
	Combinatoria	13(12,87%)
	Media	1(0,99%)
Proposiciones	Caracterización de variables aleatorias discreta y continua	1(0,99%)
	Propiedades de la función de probabilidad	15(14,85%)
	Caracterización de la distribución binomial	1(0,99%)
	La probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento	25(24,75%)
	Regla del producto	1(0,99%)
	Regla de Laplace	25(24,75%)

	La media es un representante de los datos a partir de los que ha sido calculada y no tiene por qué ser igual a uno de los valores de los datos	1(0,99%)
Procedimientos	Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial	1(0,99%)
	Cálculo de la probabilidad (de éxito) mediante la regla de Laplace	25(24,75%)
	Cálculo de probabilidad compuesta con la regla del producto	1(0,99%)
	Cálculo de combinaciones	13(12,87%)
	Cálculo de la media de la binomial mediante su expresión algebraica	1(0,99%)
Argumentos	Verbal-deductivo	2(1,98%)

Nota. Elaboración propia.

Los conceptos de variable aleatoria y distribución binomial fueron aplicados por alrededor del 1% de los jóvenes. Sin embargo conceptos vinculados a estas, como sucesos aleatorios independientes (24,75%) y combinatoria (12,87%), han sido utilizados en mayor proporción. Las proposiciones relacionadas a las tareas fueron empleadas implícitamente por los participantes. Ellos mayormente usaron la proposición que indica que la probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento (24,75%) y la regla de Laplace (24,75%), en comparación con la proposición sobre las características de la distribución binomial (0,99%), aquella en torno a la media (0,99%) y la regla del producto (0,99%).

Respecto a los procedimientos necesarios para solucionar las tareas propuestas, inicialmente el 24,75% de los jóvenes determinaron la probabilidad de elegir la respuesta correctamente de un ítem individual mediante la regla de Laplace. Luego los procedimientos vinculados a las dos estrategias para calcular las probabilidades solicitadas sólo han sido mostrados por dos estudiantes: i) uno de ellos aplicó la función de probabilidad de la binomial, además de las características de la binomial; y ii) el otro utilizó la regla del producto. También un joven calculó el valor de la media de una distribución binomial, pero sin mostrar el procedimiento de reemplazar en su expresión algebraica el valor de los parámetros n y p, ni el uso de las proposiciones involucradas (razonamiento verbal-deductivo). Por tanto, solamente dos estudiantes justificaron su respuesta mediante el razonamiento verbal-deductivo.

5.2.2.8 Resultados en el ítem 7

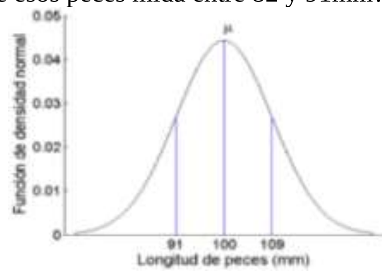
El ítem 7 presentado en la Figura 74 está integrado por: el $I_{7.2}$ extraído de González y Ojeda (2017) que valora la S-$P_{7.2}$ determinar probabilidades asociadas a una distribución normal de forma manual (Bizet et al., 2023c, p.187); el $I_{7.3}$ adaptado de MINEDUC (2019c) que evalúa la S-$P_{7.3}$ describir la tendencia de datos representados gráficamente empleando la aproximación de la binomial por la normal (Bizet et al., 2023c, p.187); y el $I_{7.4}$ tomado de Tauber (2001) que valora la

S-P$_{7.4}$ evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal (Bizet et al., 2023c, p.187).

Figura 74

Ítem 7 del cuestionario

7.2 La longitud de cierto tipo de peces sigue una distribución normal con media $\mu = 100$ y varianza $\sigma^2 = 81$. ¿Cuál es la probabilidad de que uno de esos peces mida entre 82 y 91mm?



7.3 Observa las siguientes distribuciones binomiales (Figura 1) ¿Qué sucede con el gráfico de barras cuando el n crece? Anota tu explicación

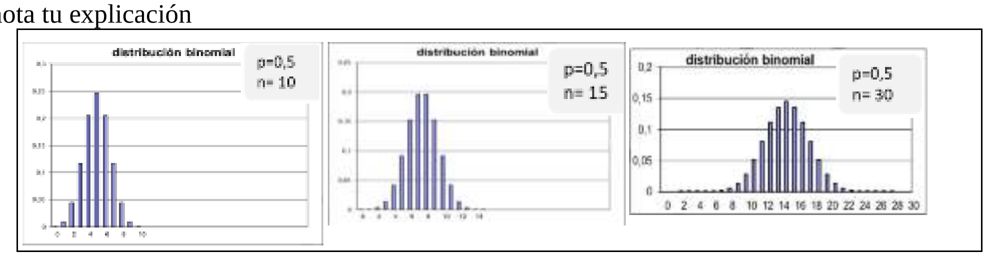


Figura 1. secuencia de distribuciones binomiales

7.4 V/F justifica tu respuesta. "en una distribución normal, el 50% de las medidas caen por encima de la media"

Nota. Extraído de González y Ojeda (2017) y Tauber (2001), y adaptado de MINEDUC (2019c).

En el I$_{7.2}$, la indagación de las respuestas entregadas por los estudiantes ha posibilitado obtener las categorías y frecuencias expuestas en el Tabla 47.

Tabla 47

Distribución de frecuencias(porcentajes) de categorías de respuestas en el I$_{7.2}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	No aplica	1(1,79%)	0	1(0,99%)
Respuesta parcialmente correcta	No aplica	0	0	0
Respuesta en blanco	No aplica	36(64,29%)	36(80%)	72(71,29%)
Respuesta errónea	E_1	2(3,57%)	1(2,22%)	3(2,97%)
	E_2	7(12,5%)	2(4,44%)	9(8,91%)
	E_3	10(17,86%)	6(13,33%)	16(15,84%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Según exhibe la Tabla 47, alrededor del 1% de los participantes respondieron correctamente (C$_2$), debido a que determinaron la probabilidad de que uno de los peces mida entre

82 y 91 milímetros, correspondiente a 0,1359 o 13,59%. Aunque ellos omitieron la argumentación de su respuesta mediante: i) el desarrollo de los procedimientos de estandarización, cálculo de probabilidades utilizando la función de densidad de la normal estándar y lectura de su tabla de probabilidades, además de la aplicación de las propiedades para el cálculo de probabilidades con distribución normal estándar; o ii) la utilización de la propiedad de los intervalos centrales ($\pm\mathbf{3\sigma}$). Ejemplo de este tipo de respuesta se muestra en la Figura 75.

Figura 75

Ejemplo de respuesta correcta en el $I_{7.2}$

Respuesta del estudiante 3 clasificada en la C_2 aproximadamente el 0.136

Nota. Elaboración propia.

Además, ninguno de los estudiantes respondió parcialmente correcto al $I_{7.3}$ (C_1), ya que no realizaron alguno de los procedimientos involucrados en la solución de la tarea, como la tipificación, cálculo de probabilidades utilizando la función de densidad de la normal estándar y/o lectura de su tabla de probabilidades. Tampoco aplicaron alguna de las proposiciones vinculada a la respuesta del problema como las propiedades para el cálculo de probabilidades con distribución normal estándar o aquella de los intervalos centrales ($\pm\mathbf{3\sigma}$).

Sin embargo gran parte de la muestra (71,29%) no respondió o mencionó no saber lo que se pregunta en el $I_{7.2}$ (C_0). Aún más, alrededor del 28% de los estudiantes respondieron de forma errónea a dicho ítem: el 2,97% de los estudiantes en vez de calcular lo solicitado determinaron la probabilidad de que uno de los peces mida menos de 91 milímetros, correspondiente a 0,1585 o 15,85%, aunque sin justificar su resultado (E_1); el 8,91% de los participantes afirmó que la probabilidad requerida es igual a 0,25 o 25%, lo cual podría deberse a que suponen que las cuatro regiones en que está dividida el área bajo la curva normal propuesta poseen igual probabilidad (E_2); y el otro 15,84% de los estudiantes señalaron un valor arbitrario que no corresponde a la probabilidad solicitada (E_3). La Figura 76 ejemplifica los errores evidenciados.

Figura 76

Ejemplos de respuestas erróneas en el $I_{7.2}$

Respuesta del estudiante 17 clasificada en E_1 15,8%	Respuesta del estudiante 99 clasificada en E_2 25%	Respuesta del estudiante 8 clasificada en E_3 3%

Nota. Elaboración propia.

Respecto a los objetos matemáticos primarios involucrados en la solución del ítem 7.2, bastantes se movilizaron en una única respuesta correcta (ver Tabla 48). Solamente un participante logró determinar la probabilidad asociadas a una distribución normal de forma manual (S-$P_{7.2}$). El lenguaje empleado por el joven para entender el enunciado de la tarea dada ha sido: i) el verbal, por medio de términos como distribución normal, media, varianza, y probabilidad; ii) el simbólico, a través de los símbolos μ y $\boldsymbol{\sigma^2}$; y iii) el gráfico, mediante la curva normal propuesta que representa la función de densidad de la normal. Además el lenguaje numérico fue utilizado por el estudiante para cuantificar la posibilidad de que uno de los peces mida entre 82 y 91 milímetros.

Tabla 48

Frecuencia de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 7.2 (n=1)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situación-Problema	S-$P_{7.2}$ Determinar probabilidades asociadas a una distribución normal de forma manual	1(0,99%)
Lenguaje	Verbal	1(0,99%)
	Numérico	1(0,99%)
	Simbólico	1(0,99%)
	Gráfico	1(0,99%)
	Tabular	0
Conceptos	Intervalo	1(0,99%)
	Variable aleatoria	1(0,99%)
	Enfoque de probabilidad clásico	1(0,99%)
	Media	1(0,99%)
	Varianza	1(0,99%)
	Desviación estándar	1(0,99%)
	Distribución normal	1(0,99%)
	Función de densidad de la normal	1(0,99%)
	Simetría	0
	Distribución normal estándar	0
Proposiciones	Caracterización de variables aleatorias discreta y continua	1(0,99%)
	Caracterización de la distribución normal	1(0,99%)
	Propiedad de los intervalos centrales (±3σ)	0
	Propiedades para el cálculo de probabilidades con distribución normal estándar	0
Procedimientos	Cálculo de probabilidad empleando la propiedad ±3σ	0
	Tipificación	0
	Cálculo de probabilidades utilizando la función de densidad de la normal estándar	0
	Lectura de tabla de la función de densidad de la normal estándar	0
Argumentos	Verbal-deductivo	0

Nota. Elaboración propia.

Los conceptos de variable aleatoria, distribución normal, función de densidad de la normal, distribución normal estándar entre otros, han sido empleados implícitamente por el estudiante. Los procedimientos vinculados a las dos estrategias para calcular la probabilidad solicitada no han sido mostrados por el estudiante: i) este no realizó la tipificación de la variable aleatoria continua y la lectura de la tabla de la distribución normal estándar, ni utilizó las propiedades para calcular probabilidades con distribución normal estándar; o ii) él no aplicó la propiedad de los intervalos centrales.

Respecto a la proposiciones, el joven empleó implícitamente la relativa a diferenciar entre variables aleatorias discretas y continuas y aquella sobre las características de la distribución normal, aunque luego no explicitó el uso de las propiedades para calcular probabilidades con distribución normal estándar o la propiedad de los intervalos centrales. Por tanto el participante no logró validar su respuesta a través del razonamiento verbal-deductivo.

Por otra parte, el análisis de las respuestas al $I_{7.3}$ permitió determinar las categorías y frecuencias exibidas en la Tabla 49.

Tabla 49

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{7.3}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	No aplica	0	0	0
Respuesta parcialmente correcta	No aplica	4(7,14%)	0	4(3,96%)
Respuesta en blanco	No aplica	22(39,29%)	30(66,67%)	52(51,49%)
Respuesta errónea	E_1	2(3,57%)	2(4,44%)	4(3,96%)
	E_2	3(5,36%)	0	3(2,97%)
	E_3	25(44,64%)	13(28,89%)	38(37,62%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como se aprecia en la Tabla 49, ninguno de los participantes respondió correctamente en el $I_{7.3}$ (C_2), ya que no lograron establer a la vez que: i) a medida que crece el n número de veces que se repite el experimento aleatorio, el gráfico de barras de la distribución binomial (con p=0,5 probabilidad de éxito en un intento) se acerca a la curva en forma de campana de la distribución normal; y ii) $\mathbf{B}(\mathbf{30};\mathbf{0},\mathbf{5})\sim\mathbf{N}(\mathbf{15};\mathbf{2},\mathbf{7})$ debido a que se cumplen las condiciones para aproximar una distribución binomial a una normal ($\mathbf{n}=\mathbf{30}\geq\mathbf{30}$, $\mathbf{n}\cdot\mathbf{p}=\mathbf{15}>\mathbf{5}$ y $\mathbf{n}\cdot(\mathbf{1}-\mathbf{p})=\mathbf{15}>\mathbf{5}$).

También alrededor del 4% de los estudiantes respondieron parcialmente correcto al $I_{7.4}$ (C_1), pues reconocieron que a medida que crece el n número de veces que se repite el experimento aleatorio, el gráfico de barras de la distribución binomial se acerca a la curva en forma de campana de la distribución normal. Ejemplos de la presente categoría se muestran en la Figura 77.

Figura 77

Ejemplos de respuestas parcialmente correctas en el $I_{7.3}$

Respuesta del estudiante 17 clasificada en la C_1 Se encuentra un comportamiento que sigue la distribución normal Respuesta del estudiante 25 clasificada en la C_1 tiende a distribuir normal	Respuesta del estudiante 26 clasificada en la C_1 Se va formando la curva de la distribución normal Respuesta del estudiante 28 clasificada en la C_1 el n crece y la campana de Gauss se hace mas y mas aparente

Nota. Elaboración propia.

Por el contrario, más de la mitad de la muestra (51,49%) no respondió o señaló no saber lo que se pregunta en el $I_{7.3}$ (C_0). Además, alrededor del 45% de los participantes respondieron incorrectamente: el 3,96% de los estudiantes afirmaron que cuando crece el n número de intentos, se mantiene el valor de p probababilidad de éxito (E_1); y el 2,97% de los participantes sostiene que cuando aumenta el n número de intentos disminuye la probabilidad de los valores de la variable aleatoria (E_2); y el 37,62% de los jóvenes restantes propusieron ideas que no tienen sentido con el contexto y tarea dada (E_3). La Figura 78 ejemplifica los errores observados.

Figura 78

Ejemplos de respuestas erróneas en el $I_{7.3}$

Respuesta del estudiante 1 clasificada en E_1 pero p se mantiene en 0.5	Respuesta del estudiante 19 clasificada en E_2 los valores del eje y disminuyen	Respuesta del estudiante 3 clasificada en E_3 la muestra se hace mas fidedigna

Nota. Elaboración propia.

Con relación a los objetos matemáticos primarios movilizados en la solución del ítem 7.3, bastantes se manifestaron en las respuestas parcialmente correctas de los participantes, aunque en una baja proporción (ver Tabla 50). Alrededor del 4% de los estudiantes lograron describir la tendencia de los datos representados en una gráfica utilizando la aproximación de la binomial por la normal (S-$P_{7.3}$).

El lenguaje utilizado por los estudiantes para entender el enunciado de la situación-problema propuesta fue: i) el verbal (3,96%), mediante términos como distribución binomial y gráfico de barras; ii) el simbólico (3,96%), a través de los símbolos n y p; y iii) el gráfico (3,96%), por medio de la secuencia de gráficos de barras expuesta que representa la función de probabilidad de la binomial B(10; 0,5), B(15; 0,5) y B(30; 0,5) respectivamente. Además el lenguaje verbal fue empleado por los participantes cuando respondieron a la tarea dada, mediante términos y expresiones verbales como: distribucion normal, campana de Gauss, curva normal, se forma, se hace aparente, tiende a y sigue un comportamiento.

Tabla 50

Frecuencia de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 7.3(n=4)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situación-Problema	S-$P_{7.3}$ Describir la tendencia de datos representados gráficamente empleando la aproximación de la binomial por la normal	4(3,96%)
Lenguaje	Verbal	4(3,96%)
	Simbólico	4(3,96%)
	Gráfico	4(3,96%)
Conceptos	Experimento aleatorio	4(3,96%)
	Variable aleatoria	4(3,96%)
	Distribución binomial	4(3,96%)
	Media	0
	Desviación estándar	0
	Distribución normal	4(3,96%)
Proposiciones	Caracterización de variables aleatorias discreta y continua	4(3,96%)
	Caracterización de la distribución normal	4(3,96%)
	Condición para aproximar una distribución binomial a una normal	0
Procedimientos	Lectura e interpretación de un gráfico de barras	4(3,96%)
	Verificación de características de la distribución normal	0
	Verificación de condiciones para la aproximación a una distribución normal	0
Argumentos	Mediante representación gráfica	4(3,96%)

Nota. Elaboración propia.

Los conceptos de variable aleatoria, distribución binomial y distribucion normal han sido empleados por alrededor del 4% de los jóvenes, aunque conceptos relacionados con estas, como media y desviación estándar no fueron usados. Dos proposiciones asociadas a la tarea han sido aplicadas implícitamente por los estudiantes. Específicamente ellos utilizaron la proposición sobre diferenciar entre variables aleatorias de tipo discretas y continuas (3,96%) y aquella relacionada con las características de distribución normal (3,96%), pero no lograron emplear las condiciones para aproximar una distribución binomial a una normal.

Respecto a los procedimientos requeridos para solucionar la situación-problema dada solamente uno de estos fue desarrollado por pocos jóvenes (3,96%), quienes leyeron los gráficos de barras propuestos y extrajeron e interpretaron la información necesaria para responder a la tarea. Por tanto, pocos participantes lograron argumentar su respuesta mediante una representación gráfica (3,96%).

En cuanto al $I_{7.4}$, la indagación de las respuestas dadas por los estudiantes hizo posible establecer las categorías y frecuencias expuestas en la Tabla 51.

Tabla 51

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{7.4}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	1(1,79%)	0	1(0,99%)
	$C_{2.2}$	0	1(2,22%)	1(0,99%)
Respuesta parcialmente	$C_{1.1}$	23(41,07%)	16(35,56%)	39(38,61%)
correcta	$C_{1.2}$	3(5,36%)	0	3(2,97%)
Respuesta en blanco	No aplica	21(37,5%)	25(55,56%)	46(45,54%)
Respuesta errónea	No aplica	8(14,29%)	3(6,67%)	11(10,89%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Según exhibe la Tabla 51, alrededor del 2% de los participantes respondieron correctamente en el $I_{7.4}$, debido a que decidieron que la afirmación "en una distribución normal, el 50% de las medidas caen por encima de la media" es verdadera argumentando su respuesta mediante indicios de las propiedades de la función de densidad de la normal: (i) el 0,99% de los estudiantes mencionaron que la normal es simétrica, pero omitieron mencionar que es respecto a su media ($C_{2.1}$); (ii) el otro 0,99% de los jóvenes insinuaron que la media es, además, la mediana de la distribución normal (aquel punto que la divide en áreas iguales y la probabilidad de ser menor a ella es 50%) ($C_{2.2}$). Ejemplos de la presente categoría se muestran en la Figura 79.

Figura 79

Ejemplos de respuestas correctas en el $I_{7.4}$

Respuesta del estudiante 8 clasificada en la $C_{2.1}$ V ya que en una distribución normal es simétrica
Respuesta del estudiante 62 clasificada en la $C_{2.2}$ verdadero ya que la media se encuentra en la mitad

Nota. Elaboración propia.

Además, el 41,58% de la muestra respondió parcialmente correcto al $I_{7.4}$. El 38,61% de los participantes decidieron que la afirmación propuesta es verdadera, pero omitieron su justificación ($C_{1.1}$), y el 2,97% de los estudiantes restantes señalaron que la proposición es verdadera, aunque argumentaron incorrectamente ($C_{1.2}$). La Figura 80 ejemplifica estos tipos de respuesta.

Figura 80

Ejemplos de respuestas parcialmente correctas en el $I_{7.4}$

Respuesta del estudiante 64 clasificada en la $C_{1.1}$	Respuesta del estudiante 13 clasificada en la $C_{1.2}$
verdadero	V pues la distribución normal estándar es aquella que tiene una media de 0 y una desviación estándar de 1.

Nota. Elaboración propia.

Mientras que el alrededor del 46% de la muestra no respondió o escribió no saber sobre lo preguntado en el $I_{7.4}$. Aún más el 10,89% de los participantes respondieron de forma incorrecta, pues decidieron que la afirmación dada es falsa, omitiento su argumentación o proponiendo ideas equícovas (E). Ejemplos de esta categoría de respuesta se presentan en la Figura 81.

Figura 81

Ejemplos de respuestas erróneas en el $I_{7.4}$

Respuesta del estudiante 48 clasificada en E	Respuesta del estudiante 22 clasificada en E
Falso	F, suele ser entre un 30% a un 40%

Nota. Elaboración propia.

Respecto a los objetos matemáticos primarios vinculados a la solución del ítem 7.4, todos ellos se han manifestado en las respuestas correctas o parcialmente correctas de los participantes, pero en un bajo porcentaje (ver Tabla 52). Alrededor del 44% de los estudiantes evaluaron la veracidad de una firmación aplicando el concepto de distribución normal (S-$P_{7.4}$). El lenguaje empleado por los estudiantes para entender el enunciado de la tarea dada ha sido tanto el verbal (43,56%), por medio de términos (distribución normal, medidas y media), como el simbólico (1,98%), a través del símbolo de porcentaje (%). Además los jóvenes utilizaron el lenguaje verbal para responder a la tarea propuesta, mediante términos (distribución normal, simetría y media) y expresión verbal (mitad).

Tabla 52

Frecuencia de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 7.4 (n=44)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situación-Problema	S-$P_{7.4}$ Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal	44(43,56%)
Lenguaje	Verbal	44(43,56%)
	Simbólico	2(1,98%)
Conceptos	Media	2(1,98%)
	Distribución normal	44(43,56%)
	Función de densidad de la normal	2(1,98%)

	Simetría	1(0,99%)
	Mediana	1(0,99%)
Proposiciones	Propiedades de la función de densidad de la normal	2(1,98%)
Procedimientos	Verificación de propiedad de la función de densidad de la normal	2(1,98%)
Argumentos	Verbal-deductivo	2(1,98%)

Nota. Elaboración propia.

El concepto de distribución normal fue aplicado por alrededor del 44% de los participantes, aunque conceptos relacionados con esta, como su función de densidad (1,98%), media (1,98%), mediana (0,99%) y simetría (0,99%) han sido utilizados en menor proporción. La proposición asociada a la tarea fue empleada por sólo dos participantes (1,98%), quienes usaron alguna propiedad de la función de densidad de la normal. Aquellos estudiantes desarrollaron el procedimiento de verificar una propiedad de la normal: i) un joven usó la propiedad que la función de densidad es simétrica con respecto a su media; y ii) el otro joven utilizó que la media es, además, la mediana de la distribución normal. Por tanto, pocos participantes validaron su respuesta mediante el razonamiento verbal- deductivo (1,98%).

5.2.2.9 Resultados en el ítem 10.a

El ítem 10.a presentado en la Figura 82, ha sido elaborado por el autor y valora la S-P$_{5.6}$ identificar la media y desviación estándar de una variable aleatoria continua.

Figura 82

Ítem 10.a del cuestionario

10. Lee el siguiente titular, luego responda la pregunta propuesta

Siempre el 16% obtiene más de 600 puntos en la PSU: DEMRE explica cómo funciona la prueba

a) Explique cómo funciona la prueba psu **solamente** mencionando los datos representados en el gráfico.

Nota. Elaboración propia.

En el I$_{10.a}$, el análisis de las respuestas proporcionadas por los estudiantes permitió obtener las categorías y frecuencias expuestas en la Tabla 53.

Tabla 53

Distribución de frecuencias(porcentajes) de categorías de respuestas en el $I_{10.a}$

Categorías	Subcategorías	Frecuencia(Porcentaje)		Total
		Grupo1	Grupo 2	
Respuesta correcta	$C_{2.1}$	1(1,79%)	0	1(0,99%)
	$C_{2.2}$	0	0	0
Respuesta parcialmente	$C_{1.1}$	2(3,57%)	2(4,44%)	4(3,96%)
correcta	$C_{1.2}$	1(1,79%)	0	1(0,99%)
Respuesta en blanco	No aplica	33(58,93%)	28(62,22%)	61(60,4%)
Respuesta errónea	E_1	2(3,57%)	6(13,33%)	8(7,92%)
	E_2	17(30,36%)	9(20%)	26(25,74%)
Total		56(100%)	45(100%)	101(100%)

Nota. Elaboración propia.

Como expone la Tabla 53, el 1% de la muestra respondió correctamente en el $I_{10.a}$, pues identificó que los puntajes obtenidos por los individuos en la prueba PSU son normalizados, donde se utiliza una escala de puntaje estándar con media de 500 puntos y desviación estándar de 110 puntos ($C_{2.1}$). La Figura 83 ejemplifica aquel tipo de respuesta. Sin embargo, ninguno de los estudiantes logró reconocer, además de la idea anterior, que: i) entre los 390 y 610 puntos hay un 68,26% de individuos, ii) entre 280 y 720 hay un 95,44%, y iii) entre 170 y 830 puntos se sitúa el 99,72% de la población ($C_{2.2}$).

Figura 83

Ejemplo de respuesta correcta en el $I_{10.a}$

Respuesta del estudiante 19 clasificada en la $C_{2.1}$

puntaje corregido es equiparado entre las distintas formas de prueba mediante un modelo estadístico que estima la habilidad de los examinados en una única escala que luego se normaliza, obteniendo como resultado un puntaje estándar (PS) con una escala con media 500 y desviación estándar 110, del grupo que rindió la prueba.

Nota. Elaboración propia.

También alrededor del 5% de la muestra respondió parcialmente correcto al $I_{10.a}$, el 3,96% de los participantes reconocieron que los puntajes en la prueba PSU son normalizados y se emplea una escala de puntaje estándar con media de 500 puntos ($C_{1.1}$), y el 0,99% de los estudiantes identificaron que en dicha prueba hay un mayor porcentaje de individuos que obtiene entre 390 y 610 puntos ($C_{1.2}$). Ejemplos de esta categoría de respuesta se exponen en la Figura 84.

Figura 84

Ejemplos de respuestas parcialmente correctas en el $I_{10.a}$

Respuesta del estudiante 26 clasificada en la $C_{1.1}$ Funciona como una distribución normal por lo que la media es 500 puntos aproximadamente
Respuesta del estudiante 16 clasificada en la $C_{1.2}$ Significa que la mayoría de los puntajes está entre 390 y 610 pts

Nota. Elaboración propia.

Por el contrario, más de la mitad de la muestra (60,4%) no respondió o señaló no saber lo que se pregunta en el $I_{10.a}$ (C_0). Además, alrededor del 34% de los estudiantes respondieron de manera errónea a aquel ítem: el 7,92% de los participantes argumentaron su respuesta a partir de creencias personales sobre el contexto de la tarea (E_1), y el otro 25,74% de los jóvenes realizaron la traducción equivocada de algunos datos del gráfico (E_2). La Figura 85 ejemplifica los errores observados.

Figura 85

Ejemplos de respuestas erróneas en el $I_{10.a}$

Respuesta del estudiante 22 clasificada en E_1 es una prueba que evalúa el conocimiento impartido durante toda tu enseñanza media y parte de la básica, se ve el conocimiento de mate-lengua-ciencias-historia. Teniendo conocimiento de esto, logramos admirar que en su mayoría de la población estudiantil no tiene los resultados esperados por diversos factores como internos (falta de estudio) o externos (falta de conocimiento de parte del ministerio en la entrega a las escuelas publicas)dejando a su mayoría con un puntaje inferior o igual a 500pts y ver si eso les sirve para la carrera que desean .
Respuesta del estudiante 44 clasificada en la E_1 la PSU funciona con puntos en cada una de sus preguntas, si se responde de manera correcta tienes más posibilidades de tener un mayor puntaje y así poder tener el adecuado para ingresar a la universidad que quieras,
Respuesta del estudiante 17 clasificada en la E_2 La PSU está diseñada para que los resultados sigan una distribución normal, es decir, se busca que alrededor del 68% de los resultados estén entro los rangos de 450 a 560 puntos aprox, mientras que el 16% tendra sobre 600 y el otro 16% bajo 450 puntos

Nota. Elaboración propia

En relación con los objetos matemáticos primarios asociados a la solución del ítem 10.a, todos ellos se han movilizado en las respuestas correctas o parcialmente correctas de los participantes, aunque en una baja proporción (ver Tabla 54). Alrededor del 1% de los estudiantes logró identificar la media y desviación estándar de una variable aleatoria continua (S-$P_{5.6}$).

El lenguaje utilizado por los jóvenes para entender el enunciado de la situación-problema propuesta fue: i) el verbal (5,94%), mediante términos y expresiones verbales como gráfico, curva, datos y puntajes normalizados; ii) el simbólico (5,94%), a través del símbolo de porcentaje (%), y

iii) el gráfico (5,94%), por medio la curva normal dada que representa la función de densidad de una normal. También el lenguaje verbal fue empleado por los participantes cuando respondieron la tarea, a través de términos y expresiones verbales como: modelo, puntajes estandarizados, media 500 puntos, desviación estándar 100 puntos, distribución normal, la mayoría de puntajes están entre 390 y 610 puntos.

Tabla 54

Frecuencia de objetos matemáticos involucrados en respuestas correctas y parcialmente correctas del ítem 10.a (n=6)

Tipo de objeto matemático	Objeto matemático	Frecuencia (Porcentaje)
Situación-Problema	S-$P_{5.6}$ Identificar la media y desviación estándar de una variable aleatoria continua	1(0,99%)
Lenguaje	Verbal	6(5,94%)
	Simbólico	6(5,94%)
	Gráfico	6(5,94%)
Conceptos	Variable aleatoria	6(5,94%)
	Intervalo	1(0,99%)
	Media	5(4,95%)
	Desviación estándar	1(0,99%)
	Distribución normal	6(5,94%)
	Función de densidad de la normal	6(5,94%)
Proposiciones	Caracterización de variables aleatorias discreta y continua	6(5,94%)
	Propiedad intervalos centrales ($\pm 3\sigma$)	1(0,99%)
Procedimientos	Lectura e interpretación de una curva normal	6(5,94%)
Argumentos	Mediante representación gráfica	6(5,94%)

Nota. Elaboración propia.

Los conceptos de distribución normal y variable aleatoria fueron empleados por alrededor del 6% de los estudiantes, aunque conceptos relacionados con ellas, como media (4,95%), desviación estándar (0,99%) e intervalo (0,99%), han sido usados en menor porcentaje. Las proposiciones relacionadas a la tarea fueron aplicadas implícitamente por los participantes, quienes principalmente usaron la proposición sobre diferenciar entre variables aleatorias de tipo discretas y continuas (5,94%), en comparación con la propiedad de los intervalos centrales (0,99%). Respecto a los procedimientos necesarios para solucionar la situación-problema, solamente el 5,94% de los jóvenes lograron leer la curva normal propuesta y extraer la información requerida. En consecuencia pocos participantes validaron su respuesta mediante una representación gráfica (5,94%).

5.2.3 Conflictos semióticos detectados en el estudio

En la presente sección son descritos los conflictos semióticos (C-S) evidenciados principalmente en las respuestas parcialmente correctas y erróneas proporcionadas por los participantes al cuestionario. La Tabla 55 expone los 17 conflictos observados en los ítems que valoran el tema de variable aleatoria.

Tabla 55

Conflictos semióticos observados en los ítems sobre variable aleatoria

Conflicto semiótico	Ítem	Frecuencia (Porcentaje)
C-S_1 Confundir una variable aleatoria con un experimento aleatorio	$I_{1.1}$	4(3,96%)
C-S_2 Confundir el recorrido con el dominio de una variable aleatoria	$I_{1.3}$	7(6,93%)
C-S_3 Confundir una variable aleatoria discreta con su función de probabilidad	$I_{1.3}$	1(0,99%)
C-S_4 Suponer que una distribución de frecuencias es uniforme	$I_{2.1a}$	8(7,92%)
C-S_5 Asumir equiprobabilidad de sucesos aleatorios compuestos	$I_{2.1b}$	7(6,93%)
C-S_6 Desconocer propiedad de la función de probabilidad "la suma de las probabilidades asignadas a todos los valores de la variable aleatoria es igual a 1"	$I_{2.1b}$	20(19,8%)
C-S_7 Desconocer la operación combinatoria "combinaciones"	$I_{2.2a}$	84(83,17%)
	$I_{2.2b}$	30(29,7%)
C-S_8 Confundir la probabilidad con el número de casos favorables de un suceso	$I_{2.2b}$	1(0,99%)
C-S_9 Asumir una relación de linealidad entre la probabilidad de un suceso (P) y parámetros (n y p)	$I_{2.2b}$	5(4,95%)
C-S_{10} Confundir la función de probabilidad con la función de densidad (o viceversa)	$I_{2.2c}$	5(4,95%)
	$I_{3.1}$	2(1,98%)
C-S_{11} Omitir título de los ejes de un gráfico	$I_{2.2c}$	4(3,96%)
	$I_{3.1}$	14(13,86%)
	$I_{4.b}$	3(2,97%)
C-S_{12} Afirmar que la función de distribución de una variable aleatoria discreta se representa gráficamente mediante un gráfico de barras	$I_{4.b}$	1(0,99%)
C-S_{13} Calcular la varianza presentando errores en su expresión algebraica	$I_{5.2b}$	4(3,96%)
C-S_{14} Confundir la varianza con la moda de una variable aleatoria discreta	$I_{5.2b}$	2(1,98%)
C-S_{15} Confundir la media con la moda de una variable aleatoria discreta	$I_{5.2c}$	9(8,91%)
C-S_{16} Argumentar una respuesta a partir de creencias personales sobre el contexto de la tarea	$I_{10.a}$	8(7,92%)
C-S_{17} Traducir del lenguaje gráfico al verbal incorrectamente	$I_{10.a}$	26(25,74%)

Nota. Elaboración propia.

El primer conflicto semiótico (C-S_1) sobre confundir una variable aleatoria con un experimento aleatorio, fue evidenciado solamente en el $I_{1.1}$ en un 3,96% de los estudiantes. El segundo conflicto (C-S_2) relativo a confundir el recorrido de una variable aleatoria con su dominio, previamente considerado como una dificultad por Miller (1998), ha sido constatado únicamente en el $I_{1.3}$ en un 6,93% de los participantes. El tercer conflicto (C-S_3) que indica confundir una variable aleatoria (discreta) con su función de probabilidad, solo fue observado en el $I_{1.3}$ en un 0,99% de la muestra.

El cuarto conflicto semiótico (C-S_4), correspondiente a suponer que una distribución de frecuencias es uniforme, es decir, afirmar que todo los valores de una variable poseen igual frecuencia absoluta, anteriormente detectado por Flores et al. (2014), fue evidenciado únicamente en el $I_{2.1a}$ en un 7,92% de los participantes. El quinto conflicto (C-S_5) relacionado con asumir equiprobabilidad de sucesos aleatorios compuestos, argumento en el cual se manifiesta el sesgo de equiprobabilidad, previamente identificado por Ruiz (2013), ha sido constatado en el $I_{2.1b}$ en un 6,93% de la muestra. El sexto conflicto (C-S_6) denota el desconocer una propiedad de la función de probabilidad, específicamente "la suma de las probabilidades asignadas a todos los valores de la variable aleatoria es igual a 1, anteriormente considerado como una dificultad por García y Sanchez (2013), fue observado sólo en el $I_{2.1b}$ en un 19,8% de los estudiantes.

El séptimo conflicto semiótico (C-S_7) relativo a desconocer la operación combinatoria involucrada, específicamente combinaciones, previamente considerado como una dificultad por Roa (2000), ha sido constatado en el $I_{2.2a}$ e $I_{2.2b}$ en un 83,17% y 29,7% respectivamente. El octavo conflicto (C-S_8) sobre confundir la probabilidad con el número de casos favorables de un suceso, anteriormente detectado por Calderón et al. (2022), fue evidenciado solamente en el $I_{2.2b}$ en un 0,99% de la muestra. El noveno conflicto (C-S_9) correspondiente a asumir una relación de linealidad entre la probabilidad de un suceso (P) y parámetros (n: número de ensayos y p: probabilidad de éxito), argumento en el cual se manifiesta el sesgo de linealidad, previamente ha sido identificado por Van Dooren et al. (2003) y fue observado en el $I_{2.2b}$ en un 4,95% de los participantes.

El décimo conflicto semiótico (C-S_{10}) que indica confundir la función de probabilidad de una variable aleatoria discreta con la función de densidad de una variable aleatoria continua (o viceversa), puntualmente cuando se debe construir el gráfico de una de ellas, ha sido evidenciado en el $I_{2.2c}$ e $I_{3.1}$ en un 4,95% y 1,98% respectivamente. El undécimo conflicto (C-S_{11}) sobre omitir título de los ejes de un gráfico, anteriormente detectado por Díaz (2018), fue constatado en: el $I_{2.2c}$ en un 3,96% de los estudiantes; el $I_{3.1}$ en un 13,86%; y el $I_{4.b}$ en un 2,97%. El duodécimo conflicto (C-S_{12}) correspondiente a afirmar que la función de distribución de una variable aleatoria discreta se representa gráficamente mediante un gráfico de barras, sólo ha sido observado en el $I_{4.b}$ en un 0,99% de la muestra.

El decimotercer conflicto semiótico (C-S_{13}) relacionado con calcular la varianza presentando errores en su expresión algebraica (ejemplo $\mathbf{Var}(\mathbf{X}) = \mathbf{E}(\mathbf{X^2}) - \mathbf{E}(\mathbf{X})$), previamente considerado por Del Pino (2019), únicamente fue constatado en el $I_{5.2b}$ en un 3,96% de los estudiantes. El decimocuarto conflicto (C-S_{14}) indica confundir la varianza con la moda de una variable aleatoria (discreta), sólo ha sido observado en el $I_{5.2b}$ en un 1,98% de la muestra. El decimoquinto conflicto (C-S_{15}) sobre confundir la media con la moda, anteriormente detectado como error por Cobo (2003), solamente fue evidenciado en el $I_{5.2c}$ en un 8,91% de los participantes.

El decimosexto conflicto semiótico (C-S_{16}) relativo a argumentar una respuesta a partir de creencias personales sobre el contexto de la tarea en vez de utilizar los datos representado en un gráfico, sólo fue observada en el $I_{10.a}$ en un 7,92% de los estudiantes. El decimoseptimo conficto (C-S_{17}) sobre traducir del lenguaje gráfico al verbal incorrectamente, previamente considerado por Berrío et al. (2022), únicamente ha sido constatado en el $I_{10.a}$ en un 25,74% de la muestra.

Por otra parte, la Tabla 56 presenta los seis conflictos semióticos evidenciados en los ítems que evalúan las distribuciones binomial y normal. En torno a la binomial, fue observado el quinto conflicto (C-S_5) relacionado con asumir equiprobabilidad de sucesos aleatorios compuestos, tanto en el $I_{6.c}$, en un 15,84% de los participantes, como en el $I_{6.d}$, en un 6,93% de los estudiantes. También ha sido constatado el noveno conflicto (C-S_9) correspondiente a asumir una relación de linealidad entre la probabilidad de un suceso (P) y parámetros (n: número de ensayos y p: probabilidad de éxito), en el $I_{6.c}$ en un 0,99% de la muestra. Aún más el decimoctavo conflicto (C-S_{18}) que indica confundir la media con la mediana, anteriormente detectado por Mayén (2009), solamente fue evidenciado en el $I_{6.e}$ en un 1,98% de los estudiantes.

Tabla 56

Conflictos semióticos observados en los ítems sobre distribuciones binomial y normal

Conflicto semiótico	Ítem	Frecuencia (Porcentaje)
C-S_5 Asumir equiprobabilidad de sucesos aleatorios compuestos	$I_{6.c}$	16(15,84%)
	$I_{6.d}$	7(6,93%)
C-S_9 Asumir una relación de linealidad entre la probabilidad de un suceso (P) y parámetros (n y p)	$I_{6.c}$	1(0,99%)
C-S_{18} Confundir la media con la mediana de una variable aleatoria discreta	$I_{6.e}$	2(1,98%)

C-S_{19} Traducir del lenguaje verbal al simbólico incorrectamente en el cálculo de probabilidades	$I_{7.2}$	3(2,97%)
C-S_{20} Desconocer las condiciones para aproximar una distribución binomial a una normal	$I_{7.3}$	49(48,5%)
C-S_{21} Desconocer la propiedad de simetría de la distribución normal	$I_{7.4}$	54(53,47%)

Nota. Elaboración propia.

En el contexto de la normal, el decimonoveno conflicto semiótico C-S_{19} sobre traducir del lenguaje verbal al simbólico incorrectamente en el cálculo de probabilidades, únicamente ha sido constatado en el $I_{7.2}$ en un 2,97% de los participantes. El vigésimo conflicto (C-S_{20}) relativo a desconocer las condiciones para aproximar una distribución binomial a una normal, anteriormente detectado por Noguera (2017) como una dificultad, sólo fue observado en el $I_{7.3}$ en un 48,5% de la muestra. Finalmente, el vigesimoprimer conflicto (C-S_{21}) indica desconocer una propiedad de la función de densidad de la normal, específicamente, que esta es es simétrica respecto del eje vertical que pasa por su media, previamente considerada por González y Ojeda (2017) como una dificultad, solamente ha sido evidenciada en el $I_{7.4}$ en un 53,47% de los participantes.

5.2.4 Significado personal logrado sobre variable aleatoria y distribuciones binomial y normal

Las respuestas de los participantes (correctas y parcialmente correctas) al cuestionario han permitido aproximarse tanto al significado personal logrado de variable aleatoria como al significado personal logra de distribuciones binomial y normal. Además se ha podido evidenciar si los objetos matemáticos primarios propuestos en el instrumento de evaluación en torno a los temas en cuestión son manipulados por los participantes.

Respecto al significado personal logrado de variable aleatoria, la Tabla 57 muestra que la mayoría de las situaciones-problemas valoradas en el cuestionario han sido resueltas por los participantes, pero en un bajo porcentaje. Las tareas en las que se evidenciaron mayor proporción de respuestas parciales o correctas han sido: i) diferenciar entre variables aleatorias y variables con dependencia funcional (S-$P_{1.1}$) con alrededor del 43% de los participantes; ii) determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde el enfoque de probabilidad clásico (S-$P_{2.2}$), con un 17,82%; y iii) representar en lenguaje gráfico la función de densidad de una variable aleatoria continua (S-$P_{3.1}$) con el 17,82%.

Tabla 57

Situaciones-problemas sobre variable aleatoria logradas por los participantes en el cuestionario

Situaciones-Problemas	Ítems						
	1	2.1	2.2	3.1	4	5	10.a
S-P$_{1.1}$ Diferenciar entre variables aleatorias y variables con dependencia funcional	42,57%						
S-P$_{1.2}$ Definir variables aleatorias finitas en experimentos aleatorios	4,95%						
S-P$_{1.4}$ Identificar recorrido o valores de una variable aleatoria finita	16,83%						
S-P$_{2.1}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial		10,89%					
S-P$_{2.2}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico			17,82%				
S-P$_{2.3}$ Definir la función de probabilidad de una variable aleatoria utilizando la terminología $P(X = x_i)$			11,88%				
S-P$_{2.4}$ Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta		15,84%					
S-P$_{2.5}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta			2,97%				
S-P$_{3.1}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua				17,82%			
S-P$_{4.1}$ Determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta					2,97%		
S-P$_{4.2}$ Definir la función de distribución de una variable aleatoria discreta					1,98%		
S-P$_{4.3}$ Representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta					0,99%		
S-P $_{5.2}$ Interpretar la media o valor esperado de una variable aleatoria discreta						1,98%	
S-P $_{5.3}$ Calcular la varianza de una variable aleatoria discreta						3,96%	
S-P$_{5.5}$ Interpretar la desviación estándar de una variable aleatoria discreta						0%	
S-P$_{5.6}$ Identificar la media y desviación estándar de una variable aleatoria continua							0,99%

Nota. Elaboración propia.

En tanto que las tareas donde se obtuvieron menor porcentaje de respuestas parcialmente correctas o correctas fueron: i) representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta (S-P$_{4.3}$) con alrededor del 1% de los participantes; ii) identificar la media y desviación estándar de una variable aleatoria continua (S-P$_{5.6}$), con aproximadamente el 1%; y iii) interpretar la desviación estándar de una variable aleatoria discreta (S-P$_{5.5}$) donde ningún estudiante logró solucionarla.

En relación con los tipos de lenguaje, la Tabla 58 expone que el lenguje más utilizado por los estudiantes en cada uno de los ítems ha sido el verbal, mediante expresiones verbales y términos. Mientras que el tipo de lenguaje menos empleado por los participantes fue: ii) el simbólico, tanto en el ítem 2.1(1,98%) como en el ítem 4 (2,97%), producto de la gran cantidad de símbolos vinculados a la variable aleatoria y que los estudiantes prefirieron expresarlos en lenguaje verbal; ii) el gráfico en el ítem 2.2 (3,96%), debido a la diversidad de elementos que intervienen en la construcción de un gráfico de barras, como por ejemplo, el título, los ejes, la escala proporcional en el eje Y, las barras de igual ancho y altura proporcional a la probabilidad correspondiente, etc.; y iii) el tabular en el ítem 1(0,99%), producto de los diferentes elementos que intervienen en la construcción de una tabla de doble entrada, tales como, el título, las etiquedas de columnas y etiquedas de filas, y el cuerpo de los datos (celdas interiores formadas de la intersección de columnas y filas), entre otros.

Tabla 58

Lenguajes y conceptos sobre variable aleatoria logrados por los participantes en el cuestionario

Objeto matemático	Ítems						
	1	2.1	2.2	3.1	4	5	10.a
Lenguajes							
Verbal	59,41%	36,63%	17,82%	24,75%	3,96%	7,92%	5,94%
Numérico		10,89%	17,82%		2,97%	3,96%	
Simbólico	16,83%	1,98%	11,88%	24,75%	2,97%	7,92%	5,94%
Gráfico	8,91%		3,96%	24,75%	3,96%		5,94%
Tabular	0,99%	16,83%	12,87%			7,92%	
Conceptos							
Experimento aleatorio	16,83%	15,84%	17,82%				
Espacio muestral	8,91%	15,84%	17,82%				
Suceso aleatorio	16,83%	15,84%	17,82%			7,92%	
Sucesos independientes			17,82%				
Función (dominio y recorrido)	8,91%			24,75%	3,96%		
Variable dependiente e independiente	42,57%						
Variable aleatoria	42,57%	36,63%	12,87%	24,75%	3,96%	7,92%	5,94%
Conjunto finito	16,83%	15,84%	17,82%				
Enfoque frecuencial de probabilidad		10,89%					
Enfoque clásico de probabilidad		15,84%	17,82%				
Distribución de frecuencias		16,83%					
Función de probabilidad		15,84%	12,87%		3,96%	7,92%	
Función de distribución					3,96%		
Intervalo				17,82%			0,99%
Función de densidad				17,82%			
Combinatoria			4,95%				
Media						1,98%	4,95%
Varianza						7,92%	
Desviación estándar						0,99%	0,99%
Distribución binomial			0%				
Función de probabilidad de la binomial			0%				

Distribución normal	5,94%
Función de densidad de la normal	5,94%

Nota. Elaboración propia.

Sobre los conceptos, la Tabla 58 presenta que el concepto más utilizados por los participantes en la mayoría de los ítems fue variable aleatoria, a excepción del ítem 2.2. En tanto que los conceptos menos aplicados en la solución de las tareas han sido: i) la distribución binomial (0%) y la combinatoria (4,95%) en el ítem 2.2; ii) la función de distribución (3,96%) en el ítem 3.1; iii) la desviación estándar y media tanto en el ítem 5 (0,99% y 1,98% respectivamente) como en el ítem 10 (0,99%, y 4,95% respectivamente).

Respecto a las proposiciones, las menos difíciles de emplear por los estudiantes en la resolución de las tareas han sido (ver Tabla 59), las características de la variable aleatoria y variables con dependencia funcional en el ítem 1 (42,57%), y las características de las variables aleatorias discretas y continuas, tanto en el ítem 2.1 (29,7%) como en el ítem 3.1 (24,75%). Sin embargo, más de la mitad de las proposiciones involucradas en las situaciones-problemas sobre variable aleatoria fueron complejas de aplicar por los participantes, entre estas: i) la variable aleatoria está caracterizada por su función de probabilidad y las propiedades de la función de probabilidad, tanto en el ítem 2.2 (2,97% cada una) como en el ítem 4 (3,96% y 2,97% respectivamente); ii) la regla del producto (3,96%) y las características de la distribución binomial (0%) en el ítem 2.2; iii) aquella en torno a la media, en el ítem 5 (1,98%) y iv) las propiedades de los intervalos centrales, en el ítem 10 (0,99%).

Tabla 59

Proposiciones, procedimientos y argumentos sobre variable aleatoria logrados por los participantes en el cuestionario

Objeto matemático	Ítems						
	1	2.1	2.2	3.1	4	5	10.a
Proposiciones							
Caracterización de la variable aleatoria y variables con dependencia funcional	42,57%						
Caracterización de variables aleatorias discreta y continua	16,83%	29,7%	12,87%	24,75%	3,96%	7,92%	5,94%
Convergencia, al crecer el número de ensayos la frecuencia relativa se va estabilizando		10,89%					
Regla de Laplace		15,84%	17,82%				
Caracterización de la variable aleatoria mediante su función de probabilidad		15,84%	2,97%		3,96%	7,92%	
Propiedades de la función de probabilidad		18,81%	2,97%		2,97%		
Regla del producto			3,96%				

La probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento			17,82%				
Caracterización de la variable aleatoria mediante su función de distribución					3,96%		
La media es un representante de los datos a partir de los que ha sido calculada y no tiene por qué ser igual a uno de los valores de los datos						1,98%	
Caracterización de la distribución binomial			0%				
Propiedades de la función de densidad				17,82%			
Propiedad de los intervalos centrales (±3σ)							0,99%
Procedimientos							
Construcción del espacio muestral	8,91%						
Partición disjunta del espacio muestral	8,91%	15,84%	3,96%				
Construcción un diagrama	8,91%						
Reproducción de experimentos aleatorio conservando las condiciones iniciales		16,83%					
Cálculo de frecuencias absolutas		10,89%					
Cálculo de probabilidades con la regla de Laplace		15,84%	17,82%				
Cálculo de combinaciones			4,95%				
Cálculo de probabilidad compuesta con la regla del producto			3,96%				
Construcción de un gráfico de barras			2,97%				
Lectura e interpretación de un gráfico de barras					3,96%		
Construcción de la gráfica de una función real en el plano cartesiano				13,86%	0,99%		
Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de distribución					3,96%		
Lectura de una tabla de probabilidad						7,92%	
Cálculo de medida de dispersión asociada a una variable aleatoria discreta empleando su expresión algebraica						7,92%	
Lectura e interpretación de una curva normal							5,94%
Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial			0%				
Argumentos							
Verbal-deductivo	0%	10,89%	3,96%			0%	
Mediante representación gráfica	8,91%		2,97%	13,86%	3,96%		5,94%

Nota. Elaboración propia.

El procedimiento que los estudiantes mayormente desarrollaron ha sido calcular probabilidades utilizando la regla de Laplace en el ítem 2.2 (17,82%). En cambio los procedimientos que más dificultó realizar a los participantes fueron: i) calcular combinaciones (4,95%), construir un gráfico de barras (2,97%), y calcular probabilidades empleando la regla del producto (3,96%) y utilizando la función de probabilidad de la binomial (0%) en el ítem 2.2; y ii) leer e interpretar un gráfico de barras (3,96%), construir la gráfica de una función real (0,99%) y calcular probabilidades empleando la función de distribución (3,96%) en el ítem 4.

Aunque los participantes lograron respuestas correctas o parcialmente correctas en las diferentes tareas y evidenciaron la manipulación de diversos objetos matemáticos primarios, solamente ha sido posible observar en un bajo porcentaje los tipos de argumentos involucrados. En general las justificaciones más empleadas fueron el razonamiento verbal-deductivo en el ítem 2.1 con el 10,89% de los jóvenes y la argumentación por medio de la representación gráfica en el ítem 3.1 con el 13,86%. Mientras que ninguno de los estudiantes ha logrado utilizar en el ítem 1 e ítem 5 el tipo de validación primero nombrada.

Con relación al significado personal logrado de distribuciones binomial y normal, la Tabla 60 expone que todas las situaciones-problemas evaluadas en el cuestionario han sido resueltas por los participantes, aunque en una baja proporción. Las tareas en las que se evidenciaron mayor porcentaje de respuestas parciales o correctas han sido: i) determinar probabilidades asociadas a una distribución binomal de forma manual ($S\text{-}P_{6.4}$) con alrededor del 15% de los estudiantes; y ii) evaluar la veracidad de afirmaciones utilizando el concepto de distribución binomial ($S\text{-}P_{7.4}$), con un 43,56%.

Tabla 60

Situaciones-problemas sobre distribuciones binomial y normal logradas por los participantes en el cuestionario

Situaciones-Problemas		Ítems			
		6	7.2	7.3	7.4
$S\text{-}P_{6.1}$ Identificar situaciones que pueden modelizarse a través de una distribución binomial		0,99%			
$S\text{-}P_{6.2}$ Determinar los parámetros asociados a una distribución binomial como n, p y q	n	1,98%			
	p	11,88%			
	q	12,87%			
$S\text{-}P_{6.3}$ Calcular la media de una distribución binomial		0,99%			
$S\text{-}P_{6.4}$ Determinar probabilidades asociadas a una distribución binomial de forma manual		14,85%			
$S\text{-}P_{7.2}$ Determinar probabilidades asociadas a una distribución normal de forma manual			0,99%		
$S\text{-}P_{7.3}$ Describir la tendencia de datos representados gráficamente empleando la aproximación de la binomial por la normal				3,96%	
$S\text{-}P_{7.4}$ Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal					43,56%

Nota. Elaboración propia.

En tanto que las tareas donde se lograron menor porcentaje de respuestas parcialmente correctas o correctas fueron: i) identificar situaciones que pueden modelizarse mediante una distribución binomial ($S\text{-}P_{6.1}$) con el 0,99% de los estudiantes, ii) calcular la media de una distribución binomial ($S\text{-}P_{6.3}$) con alrededor del 1%; y iii) determinar probabilidades asociadas a una distribución normal de forma manual ($S\text{-}P_{7.2}$) con un 0,99%.

Respecto a los tipos de lenguaje (ver Tabla 61), en el contexto de la distribución binomial el lenguje más utilizado por los estudiantes ha sido tanto el verbal (24,75%), mediante el uso de términos y expresiones verbales, como el numérico (24,75%), empleado para cuantificar la posibilidad de ocurrencia de un suceso. Mientras que el tipo de lenguaje menos usado fue el simbólico (4,95%), debido a la diversidad de símbolos relacionado con la distribución binomial que los participantes han optado por expresarlos en lenguaje verbal. En el ámbito de la distribución normal, el tipo de lenguaje más empleado por los participantes ha sido el verbal en el ítem 7.4 (43,56%), por medio de términos y expresiones verbales. En cambio el lenguaje menos utilizado fue el tabular en el ítem 7.2 (0%), producto que ninguno de los jóvenes logró usar la tabla de distribución normal estándar para calcular probabilidades.

Tabla 61

Lenguajes y conceptos sobre distribuciones binomial y normal logrados por los participantes en el cuestionario

Objeto matemático	Ítems			
	6	7.2	7.3	7.4
Lenguajes				
Verbal	24,75%	0,99%	3,96%	43,56%
Numérico	24,75%	0,99%		
Simbólico	4,95%	0,99%	3,96%	1,98%
Gráfico		0,99%	3,96%	
Tabular		0%		
Conceptos				
Experimento aleatorio	24,75%		3,96%	
Espacio muestral	14,85%			
Suceso aleatorio	24,75%			
Sucesos independientes	24,75%			
Variable aleatoria	0,99%	0,99%	3,96%	
Enfoque clásico de probabilidad	14,85%	0,99%		
Intervalo		0,99%		
Combinatoria	12,87%			
Media	0,99%	0,99%	0%	1,98%
Mediana				0,99%
Varianza		0,99%		
Desviación estándar		0,99%	0%	
Distribución binomial	0,99%		3,96%	
Función de probabilidad de la binomial	0,99%			
Distribución normal		0,99%	3,96%	43,56%
Simetría		0%		0,99%
Función de densidad de la normal		0,99%		1,98%
Distribución normal estándar		0%		

Nota. Elaboración propia.

Sobre los conceptos aplicados en la resolución de las tareas (ver Tabla 61), en el ámbito de la distribución binomial el más empleado por los estudiantes ha sido sucesos aleatorios independientes (24,75%). Al contrario, los conceptos menos utilizados fueron su función de probabilidad, su media y la varibla aleatoria, con alrededor del 1% cada uno. En el contexto de la normal, la mayoría de los conceptos han sido empleados por los participantes en un bajo porcentaje, donde el más usado fue la definición de distribución normal en el ítem 7.4 con el 43,56%. Además ninguno de los estudiantes logró aplicar el concepto de: i) simetría y distribución normal estándar en el ítem 7.2; y ii) media y desviación estándar en el ítem 7.3.

Con relación a las proposiciones (ver Tabla 62), en el contexto de la binomial, las proposisicones menos difíciles de aplicar por los participantes para solucionar las situaciones-problemas fueron la regla de Laplace (24,75%), y aquella que indica que la probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento aleatorio (24,75%). Aunque más de la mitad de las proposiciones involucradas en las tareas han sido complejas de emplear por los estudiantes, entre ellas: i) las características de variables aleatorias discretas y continuas (0,99%); ii) la regla del producto (0,99%); iii) las características de la distribución binomial (0,99%); y iv) aquella en torno a la media (0,99%).

En el ámbito de la normal, las mitad de las proposiciones involucradas en la solución de las tareas han sido empleadas por los jóvenes en una baja proporción, estas son: i) las características de las variables aleatorias discretas y continuas, tanto en el ítem 7.2 (0,99%) como en el ítem 7.3 (3,96%); ii) las características de la distribución normal en el ítem 7.2 (0,99%) e ítem 7.3 (3,96%); y ii) las propiedades de la función de densidad de la normal en el ítem 7.4 (1,98%). Mientras que la otra mitad de las proposiciones ninguno de los participantes logró aplicarlas, entre ellas, las propiedades para el cálculo de probabilidades con distribución normal estándar y aquella relacionada con los intervalos centrales en el ítem 7.2, además de las condiciones para aproximar una distribución binomial por una normal en el ítem 7.3.

Tabla 62

Proposiciones, procedimientos y argumentos sobre distribuciones binomial y normal logrados por los participantes en el cuestionario

Objetos matemáticos	Ítems			
	6	7.2	7.3	7.4
Proposiciones				
Caracterización de variables aleatorias discreta y continua	0,99%	0,99%	3,96%	
Regla de Laplace	24,75%			
Propiedades de la función de probabilidad	14,85%			
Regla del producto	0,99%			

La probabilidad de éxito y la probabilidad de fracaso es constante en cada experimento	24,75%			
Caracterización de la distribución binomial	0,99%			
La media es un representante de los datos a partir de los que ha sido calculada y no tiene por qué ser igual a uno de los valores de los datos	0,99%			
Caracterización de la distribución normal		0,99%	3,96%	
Propiedades de la función de densidad de la normal				1,98%
Propiedad de los intervalos centrales ($\pm 3\sigma$)		0%		
Propiedades para el cálculo de probabilidades con distribución normal estándar		0%		
Condición para aproximar una distribución binomial a una normal			0%	
Procedimientos				
Cálculo de probabilidad (de éxito) con la regla de Laplace	24,75%			
Cálculo de combinaciones	12,87%			
Cálculo de probabilidad compuesta con la regla del producto	0,99%			
Lectura e interpretación de un gráfico de barras			3,96%	
Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial	0,99%			
Cálculo de la media de la binomial mediante su expresión algebraica	0,99%			
Verificación de característica o propiedades de la distribución normal			0%	1,98%
Cálculo de probabilidad empleando la propiedad $\pm 3\sigma$		0%		
Tipificación		0%		
Cálculo de probabilidades utilizando la función de densidad de la normal estándar		0%		
Lectura de tabla de la función de densidad de la distribución normal estándar		0%		
Verificación de condiciones para la aproximación a una distribución normal			0%	
Argumentos				
Verbal-deductivo	1,98%	0%		1,98%
Mediante representación gráfica			3,96%	

Nota. Elaboración propia.

Con relación a los procedimientos, en el ámbito de la distribución binomial, los procedimientos que los participantes mayormente realizaron fueron calcular probabilidades usando la regla de Laplace (24,75%) y calcular combinaciones (12,87%). En tanto que los procedimientos que más dificultó desarrollar a los estudiantes han sido calcular probabilidades con la regla del producto (0,99%) y mediante la función de probabilidad de la binomial (0,99%), además de calcular la media a través de su expresión algebraica (0,99%).

En el contexto de la distribución normal, sólo dos de los siete procedimientos involucrados en la solución de las tareas fueron desarrollados por los jóvenes, pero en una baja proporción. Aquellos son leer e interpretar un gráfico de barras en el ítem 7.3 (3,96%) y verificar propiedades o características de la normal en el ítem 7.4 (1,98%). Sin embargo, ninguno de los estudiantes ha logrado desarrollar los procedimientos: i) calcular probabilidades empleando la propiedad de los intervalos centrales o mediante la función de densidad de la normal estándar, la tipificación y leer la tabla de la distribución normal estándar en el ítem 7.2; y ii) verificar las condiciones para aproximar una distribución binomial por una normal en el ítem 7.3.

Respecto a los tipos de argumentación, en el ámbito de la binomial ha sido posible observar en una baja proporción, que los estudiantes justificaron su respuesta mediante el razonamiento verbal-deductivo (1,98%). En el contexto de la normal, los participantes utilizaron los dos tipos de argumentación involucrados en las resolución de las tareas, aunque en un bajo porcentaje. Puntualmente ellos lograron validar su respuesta por medio del uso de representación gráfica (3,96%) en el ítem 7.3 y justificar su solución a través del razonamiento-verbal deductivo (1,98%) en el ítem 7.4.

5.2.5 Síntesis de los resultados

5.2.5.1 Puntuación según ítem, campo de problema y total del cuestionario

A partir de la distribución de frecuencias de las categorías de respuestas obtenida en cada ítem se han calculado estadísticos descriptivos para cada ítem (ver Tabla 63). En las tareas sobre variable aleatoria, la muestra alcanzó la mayor puntuación media en el $I_{1.1}$ que valora la $S\text{-}P_{1.1}$ diferenciar entre variables aleatorias y variables con dependencia funcional y la $S\text{-}P_{1.2}$ definir variables aleatorias finitas en experimentos aleatorios (0,64 puntos). Mientras que la muestra obtuvo la puntuación media más baja en: i) el $I_{4.b}$ que evalúa la $S\text{-}P_{4.3}$ representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta (0,02 puntos); y ii) el $I_{5.2c}$ que evalúa la $S\text{-}P_{5.2}$ interpretar la media de una variable aleatoria discreta y la $S\text{-}P_{5.5}$ interpretar la desviación estándar de una variable aleatoria discreta (0,02 puntos).

Tabla 63

Estadísticos descriptivos en la puntuación de cada ítem

Ítem	Mínimo	Máximo	Media	Desviación estándar
1.1	0	2	0,64	0,576
1.3	0	2	0,20	0,469
2.1a	0	2	0,48	0,687
2.1b	0	2	0,37	0,717
2.2a	0	2	0,23	0,527
2.2b	0	2	0,17	0,471
2.2c	0	2	0,05	0,260
3.1	0	2	0,39	0,721
4.a	0	2	0,05	0,296
4.b	0	2	0,02	0,199
5.2b	0	2	0,12	0,431
5.2c	0	1	0,02	0,140
6.c	0	2	0,40	0,736
6.d	0	2	0,35	0,685

6.e	0	2	0,02	0,199
7.2	0	2	0,02	0,199
7.3	0	1	0,04	0,196
7.4	0	2	0,46	0,539
10.a	0	2	0,07	0,292

Nota. Elaboración propia.

En las tareas sobre distribuciones binomial y normal, la muestra alcanzó la mayor puntuación media en el $I_{6.c}$ que principalmente evalúa la S-$P_{6.4}$ determinar probabilidades asociadas a una distribución binomial de forma manual (0,4 puntos), y en el $I_{7.4}$ que valora la S-$P_{7.4}$ evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal (0,46 puntos). En cambio, la muestra consiguió la menor puntuación media en el $I_{6.e}$ que principalmente evalúa la S-$P_{6.3}$ calcular la media de una distribución binomia (0,02 puntos), y en el $I_{7.2}$ que valora la S-$P_{7.2}$ determinar probabilidades asociadas a una distribución normal de forma manual (0,02 puntos).

Respecto al puntaje máximo de cada ítem (2 puntos), este no fue logrado en el $I_{5.2c}$ que valora la S-$P_{5.2}$ interpretar la media de una variable aleatoria discreta y la S-$P_{5.5}$ interpretar la desviación estándar de una variable aleatoria discreta, debido a que ninguno de los participantes interpretaron estas dos medidas a la vez, sólo la primera señalada. También el puntaje máximo no ha sido conseguido en el $I_{7.3}$ que evalúa la S-$P_{7.3}$ describir la tendencia de datos representados gráficamente empleando la aproximación de la binomial por la normal, pues ningún estudiante comprobó las condiciones para aproximar una distribución binomial a una normal y determinó la aproximación $(B(30; 0{,}5) \sim N(15; 2{,}7))$.

Por otra parte, la puntuación de los ítems que componen cada campo de problema fueron sumadas, estableciéndose que en el C-P_1, C-P_4, C-P_5, C-P_6 y en el C-P_7 se podía lograr un máximo de 6 puntos y en el C-P_2 a lo más 8 puntos. Luego ha sido calculada la puntuación media en cada campo de problema (ver Tabla 64), evidenciándose que la muestra obtuvo una puntuación media deficiente, con valores menores a 2 en todos los campos de problemas. La mayor puntuación media se alcanzó en el C-P_1 relativo a identificar la variable aleatoria como una función (1,228 puntos) y el C-P_2 en torno a reconocer la distribución de probabilidad de una variable aleatoria discreta (1,238 puntos). Mientras que la puntuación media más baja (indica escasez de respuestas correctas) se obtuvo en el C-P_4 relacionado con establecer la función de distribución de una variable aleatoria discreta (0,119 punto) y el C-P_5 sobre utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria (0,208 puntos), presentándose una variabilidad similar en estos.

Tabla 64

Estadísticos descriptivos en la puntuación de cada campo de problema

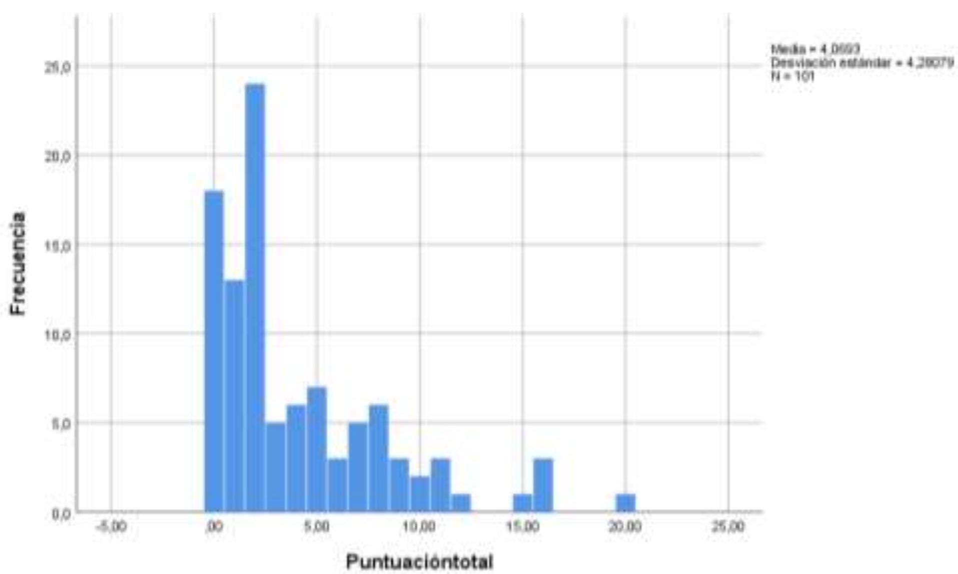
Campo de Problema	Mínimo	Máximo	Media	Desviación estándar
C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico	0	6	1,228	1,318
C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria	0	7	1,238	1,976
C-P_4 Establecer la función de distribución de una variable aleatoria discreta	0	4	0,119	0,553
C-P_5 Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria	0	4	0,208	0,668
C-P_6 La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real	0	6	0,762	1,415
C-P_7 La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real	0	3	0,515	0,626

Nota. Elaboración propia.

En relación con la puntuación total, ninguno de los participantes logró en el cuestionario el puntaje máximo (38 puntos), además según expone la Figura 86 el puntaje total varió entre 0 y 20 puntos, donde se alcanzó una media de 4 puntos. Por tanto el rendimiento de los participantes sobre la comprensión de variable aleatoria y distribuciones binomial y normal se puede calificar de deficiente.

Figura 86

Histograma de puntuación total del cuestionario



Nota. Elaboración propia.

5.2.5.2 Diferencia entre grupos de la muestra

Fueron calculados algunos estadísticos descriptivos en cada campo de problema, diferenciando entre los grupos que integran la muestra (ver Tabla 65). De esta manera es posible evidenciar en cada campo de problema, que el primer grupo (estudiantes de Pedagogía en Matemática) alcanzó mayor puntuación media que el segundo grupo (estudiantes de Pedagogía en Educación Especial). Además, G_1 logró la mayor puntuación media en el C-P_2 sobre reconocer la distribución de probabilidad de una variable aleatoria discreta (2 puntos), mientras que G_2 en el C-P_1 relativo a identificar la variable aleatoria como una función (0,444 puntos). También, ambos grupos obtuvieron la puntuación media más baja en el C-P_4 en torno a establecer la función de distribución de una variable aleatoria discreta (0,214 y 0 puntos respectivamente).

Tabla 65

Estadísticos descriptivos en la puntuación de cada campo de problema por grupo

Campo de Problema	Mínimo		Máximo		Media		Desviación estándar	
	G_1	G_2	G_1	G_2	G_1	G_2	G_1	G_2
C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico	0	0	6	2	1,857	0,444	1,407	0,586
C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria	0	0	7	3	2,000	0,289	2,328	0,661
C-P_4 Establecer la función de distribución de una variable aleatoria discreta	0	0	4	0	0,214	0,000	0,731	0,000
C-P_5 Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria	0	0	4	1	0,339	0,044	0,859	0,208
C-P_6 La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real	0	0	6	4	1,071	0,378	1,672	0,886
C-P_7 La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real	0	0	3	2	0,607	0,400	0,679	0,539

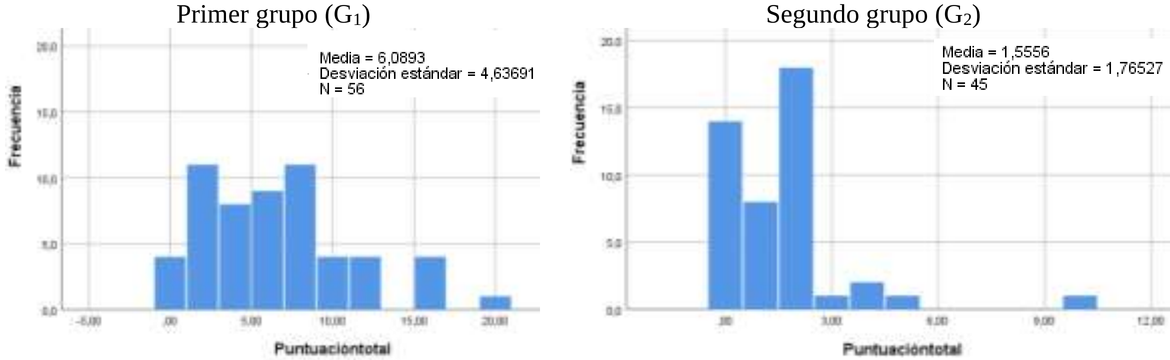
Nota. Elaboración propia.

La Figura 87 muestra que en el primer grupo la puntuación total varió entre 0 y 20 puntos, consiguiendo una media de 6 puntos, en tanto que en el segundo grupo el puntaje total varió entre 0 a 10 puntos, obteniendo una media de 1 punto. Además se observa mayor variabilidad de las puntuaciones totales en el primer grupo.

Figura 87

Distribución de la puntuación total del cuestionario por grupo

Primer grupo (G_1) Segundo grupo (G_2)



Nota. Elaboración propia.

5.2.6 Discusión de resultados

En el presente estudio, la mayoría de los objetos matemáticos involucrados en el cuestionario han sido observados en las respuestas (parcialmente correctas o correctas) de los participantes, pero en en un bajo porcentaje. Puntualmente los resultados sobre la comprensión de variable aleatoria evidencian en torno al primer campo de problema que, alrededor del 43% de los participantes lograron diferenciar entre variables aleatorias y variables con dependencia funcional, resultado más alentador que el 17% obtenido en el estudio de Pérez y Parraguez (2013). Además el 16,83% de los estudiantes establecieron el recorrido de una variable aleatoria discreta, cifra que está en consonancia con la evidenciada en la investigación de Vilca (2015).

Tambien se observó que el 17,82% de los jóvenes representaron en lenguaje gráfico una función de densidad, particularmente graficaron una función constante en un intervalo, resultado más descendido que el 56% evidenciado por Mevarech y Kramarsky (1997). Según Hausberger et al. (2021), una posible explicación de este escenario podría ser que los estudiantes no han desarrollado suficientes conexiones entre los conceptos de distribución de probabilidad (de una variable aleatoria continua) y área, pues pocos participantes emplearon: el lenguaje simbólico (24,75%) y lenguaje gráfico (24,75%) de la función de densidad, el procedimiento de construir la gráfica de una función real en el plano cartesiano (13,86%) y las propiedades de la función de densidad (17,82%).

En tanto que en el segundo campo de problema, se constató que alrededor del 18% de los participantes calcularon probabilidades asociadas a valores de una variable aleatoria discreta desde el enfoque clásico de probabilidad, resultado semejante al obtenido por Salazar (2014), quien constató que alrededor del 30% de su muestra (estudiantes universitarios) determinó aquellas probabilidades. Además alrededor del 16% de los jóvenes representaron la función de probabilidad en lenguaje tabular, porcentaje similar al constatado en la investigación de Bizet y Ramos (2022).

En el campo de problema cuatro, se observó una de las tareas sobre variable aleatoria más complejas de responder para los participantes. Sólo el 1% de los jóvenes lograron representar en lenguaje gráfico la función de distribución, es decir, graficar una función escalonada, resultado más descendido que el 50% constatado en la investigación de Vargas et al. (2016). En términos de Sharma (2006), aquel escenario podría deberse a que a los estudiantes les dificultad leer e interpretar un gráfico de barras, pues sólo el 3,96% de los participantes desarrollaron dicho procedimiento en la resolución de la tarea señalada. Asimismo, en el campo de problema cinco, se observó otra de las tareas difíciles de solucionar para los jóvenes. Ninguno de los participantes logró interpretar la desviación estándar de una variable aleatoria discreta, cifra menos alentadora que el 33,4% alcanzado en la investigación de Armah et al. (2016).

Respecto a los resultados sobre la comprensión de las distribuciones binomial y normal, estos mostraron en torno al campo de problema seis que alrededor del 1% de los participantes lograron identificar la situación modelizada mediante una binomial. Este hallazgos es similar al de Wroughton y Cole (2013), quienes constataron en una prueba sobre distribuciones de probabilidad, que la mayoría de las respuestas incorrectas de su muestra fueron cuando se debían reconocer situaciones binomiales. También, se evidenció que aproximadamente el 15% de los estudiantes calcularon probabilidades asociadas a una distribución binomial de forma manual, resultado semejante al 19% conseguido en la investigación de Toledo et al. (2019). Aún más sólo el 1% de los participantes calcularon la media de una binomial, según Taufiq et al. (2020) este escenario podría deberse a que los libros de texto abordan superficialmente aquellos conceptos o en ocasiones omiten su tratamiento.

En el campo de problema siete, se evidenció que alrededor del 44% de los estudiantes evaluaron la veracidad de una afirmación utilizando el concepto de distribución normal, porcentaje más descendido que el 69% obtenido en el estudio de Tauber (2001). Asimismo, aproximadamente el 4% de los participantes lograron utilizar la aproximación de la binomial por la normal cuando describieron la tendencia de datos representados gráficamente, hallazgo que concuerda con lo observado por Salinas-Herrera y Salinas-Hernández (2022), quienes constataron que sus estudiantes, mediante la observación de un gráfico, no pudieron intepretar que existe una relación entre la distribución binomial y la normal. Aún más, sólo el 1% de los jóvenes calcularon la probabilidad asociada a una distribución normal de forma manual, porcentaje igual de bajo que el 14% conseguido en la investigación de Bansilal (2014).

Finalmente el bajo desempeño obtenido por los participantes en las diversas tareas muestran un escenario complejo para la educación estadística. Esta realidad podría deberse a que la Unidad de Currículo y Evaluación del Ministerio de Educación de Chile, desde el 2020 hasta el 2022, puso a disposición del sistema escolar una priorización del currículo vigente (MINEDUC, 2023), por los siguientes motivos: i) contexto nacional, un estallido social ocurrido en octubre del 2019, que generó la inasistencia a la escuela hasta el término del año escolar (diciembre); y ii) contexto internacional, la emergencia sanitaria originada por la pandemia de la Covid-19.

5.2.7 Conclusiones

Los resultados de este estudio permitieron identificar qué comprenden los estudiantes chilenos egresados de educación escolar sobre variable aleatoria y distribuciones binomial y normal, desde la perspectiva del Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos.

Sobre la variable aleatoria, el significado personal logrado por los participantes carece de ciertos objetos matemáticos primarios que lo diferencian del significado institucional evaluado. Por ejemplo, la situación-problema relativa a interpretar la desviación estándar de una variable aleatoria discreta, ninguno de los jóvenes logró responderla. Además limitados estudiantes de la muestra han logrado emplear los distintos objetos matemáticos involucrados en la solución de las tareas evaluadas. En consecuencia los estudiantes egresados de educación escolar chilena poseen una comprensión baja de variable aleatoria, resultado que difiere del obtenido en la investigación de Ruiz (2013), donde estudiantes españoles de segundo año de universidad lograron una comprensión parcial del tema en cuestion.

Respecto a la distribución binomial, el significado personal logrado por los estudiantes de la muestra es coherente con el significado institucional evaluado en el cuestionario, aunque pocos participantes lograron utilizar la diversidad de objetos matemáticos primarios asociados a la resolución de las situaciones-problemas valoradas. Por tanto los participantes poseen una comprensión baja de distribución binomial, resultado que está en consonancia con el obtenido en la investigación de Landín (2013), luego que él desarrolló un proceso de instrucción sobre el tema en estudiantes escolares (17 a 18 años).

Con relación a la distribución normal, el significado personal logrado por los jóvenes participantes carece de objetos matemáticos primarios que lo diferencian del significado institucional evaluado, tales como: el lenguaje tabular; los conceptos de media, desviación estándar, simetría y distribución normal estándar; la propiedad de los intervalos centrales, aquellas para calcular probabilidades con distribución normal estándar y las condiciones para aproximar una binomial por una normal; y el proceso de calcular probabilidades, la tipificación y la lectura de la tabla de la normal estándar. También pocos estudiantes de la muestra lograron aplicar algunos objetos primarios asociados a la solución de las tareas valoradas. Por consiguiente los participantes poseen una comprensión baja de la distribución normal, hallazgo que difiere con lo observado en la investigación de Tauber (2001), donde estudiantes universitarios lograron una comprensión parcial al finalizar un proceso de instrucción en torno al tópico.

Por tanto, es importante que los Ministerios de Educación, como el chileno, consideren los resultados de esta investigación para diseñar programas de estudios que fomenten la comprensión de variable aleatoria y modelos de probabilidad mediante el uso de tecnología. Debido a que esta es una herramienta fundamental en la enseñanza-aprendizaje de la matemática y mejora el aprendizaje de los estudiantes (NCTM, 2000). Especificamente la tecnología ayuda a comprender y aplicar conceptos de probabilidad, como las distribuciones de probabilidad de una variable aleatoria (NCTM, 2000; Bargagliotti et al., 2020), la distribución binomial (NCTM, 2000) y la distribución normal (Bargagliotti et al., 2020).

Así para abordar la variable aleatoria discreta y su función de probabilidad se sugiere: i) desarrollar las actividades *suma de puntos al lanzar dos dados* y *recogida de datos de la clase* (Godino, 2013); y ii) trabajar con un software en la simulación del lanzamiento de los dos dados 100 y 1000 veces, por ejemplo usar GeoGebra (https://www.geogebra.org/m/byV5AfzR), con la finalidad de analizar la convergencia de las frecuencias relativas a las correspondientes probabilidades conforme crece el número de repeticiones de un experimento. Este tipo de prácticas en términos de Lee et al. (2010): “fomenta la comprensión de los estudiantes sobre la conexión entre las frecuencias relativas de los datos empíricos y la probabilidad derivada teóricamente, y el uso de herramientas tecnológicas en el aprendizaje de la probabilidad” (p.91).

Posteriormente, para el tratamiento de la la función de densidad se propone desarrollar *el problema de las citas* y *el problema del volcan Aso* (Derouet, 2019), donde se trabaja con histogramas mediante el uso del software GeoGebra, con el propósito de relacionar la probabilidad de variables aleatorias continuas con el concepto de área (integral definida). En tanto que para abordar la binomial se sugiere desarrollar la actividad *ilustrando la distribución binomial* mediante el juego de dados GOLO en línea (Stephenson et al., 2009). Mientras que para trabajar la normal se propone resolver el problema *exámen de acceso a la universidad de María* por medio del uso de un software (Salinas-Herrera y Salinas-Hernández, 2022).

Por otra parte, entre las limitaciones de este estudio se considera el tipo de muestra (no probabilística) y la forma de aplicación de cuestionario (en línea), debido al estado de pademia por la Covid-19 presentado durante el proceso de recolección de datos. Finalmente las evidencias obtenidas en la presente investigación plantea el desafío de diseñar e implementar una unidad didáctica sobre los tema en cuestión con uso de tecnología con el propósito de aportar a la enseñanza-aprendizaje de la probabilidad.

CAPÍTULO 6: CONCLUSIONES

6.1 Introducción

En el presente capítulo se exponen las conclusiones finales de esta tesis doctoral, organizadas en tres apartados. En el primer de ellos, es presentada una discusión sobre el logro de los objetivos e hipótesis de investigación planteados en el Capítulo 1. En el apartado 6.3 se muestra una síntesis de las principales contribuciones de la tesis para la Educación Estadística. En el último apartado, son analizadas las limitaciones de esta investigación y presentadas algunas líneas para futuros trabajos.

6.2 Conclusiones sobre los Objetivos e Hipótesis de Investigación

6.2.1 Conclusiones sobre el primer Objetivo e Hipótesis

(O1): *Desarrollar una revisión bibliográfica sobre la enseñanza y aprendizaje de variable aleatoria y distribuciones binomial y normal.*

Para alcanzar esta meta ha sido realizado un estudio presentado en el Capítulo 2. En el Estudio 1 (Bizet et al., 2022) fueron analizadas las principales investigaciones que reportó la literatura en el campo de educación matemática, entre los años 1990 y 2020, sobre la enseñanza y aprendizaje de la variable aleatoria y su distribución de probabilidad, tanto a nivel escolar como universitario. Los resultados evidenciaron que se ha encontrado una menor cantidad de estudios que abordan sus aspectos epistemológicos y didácticos (enseñanza) en comparación con los que tratan sus elementos cognitivos (aprendizaje).

Con relación al ámbito epistemológico, se aprecian investigaciones que han identificado de manera superficial elementos de la variable aleatoria (problemas, procedimientos, diferentes representaciones, etc.) necesarios para su enseñanza en la escuela (Ortiz, 2002) y universidad (Ruiz, 2013). Respecto al modelo binomial, dentro de lo que se ha buscado, no se han hallado estudios que entreguen directrices sobre los elementos principales a considerar en su enseñanza (aunque en los antecedentes complementario al Estudio 1, es decir, trabajos publicados posterior al 2020, se encontró uno sobre su origen y desarrollo histórico). Contrario a lo ocurrido con la normal, los resultados obtenidos son más alentadores, dado que se encontró detalladamente los elementos fundamentales a incluir en su proceso de enseñanza, tanto en la educación superior (Tauber, 2001) como escolar obligatoria (Valverde, 2017).

Sobre los aspectos didácticos se descubrieron similar cantidad de trabajos dirigidos a la educación universitaria y escolar. Estos se centran en propuestas para la enseñanza a nivel universitario de la variable aleatoria y el modelo normal, y a nivel escolar de la binomial. Mientras que existe una menor cantidad de propuestas de instrucción en el contexto escolar en torno a la variable aleatoria y la normal, y en el ámbito universitario sobre la binomial.

Dentro de los estudios desarrollados desde una perspectiva cognitiva, se hallaron alrededor de la misma cantidad de investigaciones que abordan la comprensión de variable aleatoria y su distribución de probabilidad en la educación escolar y universitaria. Sin embargo, en la literatura indagada existen vacíos que serían importantes investigar, por ejemplo: i) de forma articulada el nivel de comprensión de la variable aleatoria y los modelos de probabilidad al término de la educación escolar; o ii) el nivel de comprensión de la distribución binomial en la educación superior. Por otra parte, se han identificado algunas dificultades que manifiestan los estudiantes a nivel escolar y/o universitario cuando trabajan con la variable aleatoria, la distribución binomial y/o la distribución normal. Sin embargo, no se han encontrado estudios que interpreten aquellas dificultades o errores en torno a los temas de interés en términos de conflictos semióticos.

De esta manera, los resultados evidenciados en el Estudio 1 confirma la hipótesis que el análisis de investigaciones previas sobre la enseñanza y aprendizaje de variable aleatoria y/o distribuciones binomial y normal muestra que han sido temas limitadamente investigados. Con relación al análisis de las investigaciones sobre la enseñanza de variable aleatoria y modelos de probabilidad en la escuela, se reconocieron algunas investigaciones, pero las halladas no desarrollaron una indagación profunda de los elementos (situaciones-problemas, proposiciones, procedimientos, lenguajes, etc.) que intervienen en sus tareas, tampoco realizaron un análisis de los campos de problemas que pueden identificarse. Respecto a la indagación de las investigaciones en torno al aprendizaje de los temas de interés a nivel escolar, no se han hallado estudios que analizan articuladamente la comprensión de variable aleatoria y modelos de probabilidad al término de la escuela. Tampoco se encontraron estudios que interpreten las dificultades o errores sobre aquellos temas en términos de conflictos semióticos.

6.2.2 Conclusiones sobre el segundo Objetivo e Hipótesis

(O2): *Analizar el tratamiento de variable aleatoria y distribuciones binomial y normal en la normativa curricular y libros de texto chilenos.*

Este objetivo fue logrado mediante el desarrollo de dos estudios (Estudio 2 y Estudio 3) expuestos en el Capítulo 3. En el primero de ellos (Bizet et al., 2023a), se analizaron las situaciones-problemas sobre variable aleatoria y distribuciones binomial y normal en una muestra de cinco libros de texto chilenos organizadas en dos series, donde cada una abarcó los grados 10 (15 a 16 años), 11 (16 a 17 años) y 12 (17 a 18 años). Inicialmente, con base en la Guía de Situaciones-Problemas sobre Variable Aleatoria y sus aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar Chileno (GSP-VADP: Bizet et al., 2023c), fueron clasificadas las actividades presentes en los libros indagados, reconociendo el o los tipos de situaciones-problemas involucradas en cada una. Posteriormente se contrastaron las situaciones-problemas en torno a los temas en cuestión propuestas en el currículo escolar y libros de texto chilenos.

Los resultados del Estudio 2 han evidenciado respecto a la diversidad de situaciones-problemas sobre la variable aleatoria, que en los libros analizados fueron observadas 24 tipos, organizadas en cinco campos de problemas: el C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico; el C-P_2 Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria; el C-P_3 Reconocer la distribución de probabilidad de una variable aleatoria continua como una herramienta que permite ver la variación aleatoria; C-P_4 Establecer la función de distribución de una variable aleatoria discreta; y el C-P_5 Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria.

Asimismo, en los textos escolares indagados fue evidenciada gran variedad de situaciones-problemas sobre la aplicación de la variable aleatoria en distribuciones de probabilidad. Particularmente para el modelo binomial han sido observadas cuatro tipos, estructurada en un campo de problema, el C-P_6 La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real. Mientras que para el modelo normal se constaron 9 tipos, organizadas en dos campos de problemas: el C-P_7 La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real; y el C-P_8 Aproximar distribuciones de variables aleatorias discretas con un número grande de valores.

De esta manera, se observó que hay situaciones-problemas fomentadas en los libros de texto analizados que no fueron sugeridas en el currículo chileno, tales como: diferenciar entre variables aleatorias discretas y continuas, calcular el valor de incógnitas de manera que la función propuesta sea de probabilidad (para ambos tipos de variables señaladas), calcular probabilidades en una distribución normal estándar, representar en un gráfico la distribución normal y calcular los valores correspondientes a una probabilidad dada en el contexto de la distribución normal. También fue evidenciado que existen situaciones-problemas identificadas en el currículo chileno que no fueron abordadas en los textos escolares o poseen mínima presencia, como: calcular probabilidades asociadas a distribuciones binomial y normal empleando una herramienta tecnológica, y diferenciar entre variables aleatorias y variables con dependencia funcional, respectivamente. Por tanto, los resultados demuestran falta de alineación entre la propuesta del MINEDUC (2015; 2019a) y los textos analizados.

En tanto que en el Estudio 3 (Bizet et al., 2023b), se indagó el lenguaje, conceptos, proposiciones, procedimiento y argumentos vinculados a la variable aleatoria y las distribuciones binomial y normal en la muestra de cinco libros de texto chilenos empleada previamente (Estudio 2). Los objetos matemáticos de interés primero han sido reconocidos en el currículo escolar chileno, luego fueron identificados en la muestra de libros y finalmente se contrastaron los resultados obtenidos. Sobre el lenguaje relacionado con la variable aleatoria se evidenció que entre los lineamientos curriculares y las dos series de libros existe congruencia, debido a que en cada documento se han propuesto cuatro tipos de lenguaje. Para la distribución normal se constató un escenario semejante al descrito anteriormente. En el contexto de la binomial se observó desarmonía entre los documentos indagados, ya que el currículo sugirió tres tipos de lenguajes al igual que una de las series de libros y la otra serie excluyó el lenguaje gráfico.

Respecto a los conceptos, para la variable aleatoria se ha evidenciado coherencia entre los lineamientos curriculares y las dos series de libros, pues en cada documento se presentaron siete conceptos. En el ámbito de los modelos de probabilidad se observaron discrepancias entre los documentos analizados. Para la distribución binomial el currículo promovió seis conceptos, a diferencia de las dos series de libros que omitieron algunos, una de ellas omitió su función de distribución y desviación estándar, y la otra su media, varianza y desviación estándar. En la normal el currículo sugirió siete conceptos, en cambio las dos series suprimieron algunos, una de estas descartó la función de densidad de la normal estándar, la media, la desviación, la moda y la mediana, y la otra las dos últimas señaladas.

Sobre las proposiciones se identificaron ciertas discrepancias entre los documentos indagados. Específicamente en la variable aleatoria el currículo sugirió cinco proposiciones, excluyendo su caracterización mediante la función de distribución y a través de la función de densidad, mientras que una serie de libros propuso siete y la otra seis, esta última omitió la caracterización de variable aleatoria y variables con dependencia funcional. Además, para la normal los lineamientos curriculares propusieron seis proposiciones, descartando las propiedades de la función de densidad, a diferencia de cada serie que sugirió siete. Sin embargo, en el contexto de la binomial se evidenció armonía entre los documentos analizados, ya que el currículo sugirió dos proposiciones iguales que las propuestas en cada serie.

También en los procedimientos existió desarmonía entre los documentos indagados. El currículo para la variable aleatoria promovió seis procedimientos, ignorando calcular probabilidades con su función de distribución, mientras que una serie de libros propuso siete y la otra seis, excluyendo determinar probabilidades o su dominio utilizando diagrama de árbol. En la binomial, los lineamientos curriculares presentaron tres procedimientos al igual que una serie de libros, pero la otra serie descartó calcular su media, varianza y desviación estándar mediante expresiones algebraicas. Para el ámbito de la normal el currículo sugirió tres procedimientos exceptuando la corrección por continuidad, mientras que cada serie de libros cuatro.

Asimismo, en los argumentos se encontraron incongruencias entre los documentos analizados. En la variable aleatoria los lineamientos curriculares fomentaron cinco tipos, suprimiendo el razonamiento verbal-deductivo y la justificación mediante análisis-síntesis, a diferencia de las dos series de libros: una promovió cinco, descartando las simulaciones utilizando herramienta tecnológica y objeto manipulable; y la otra promovió seis, omitiendo la última nombrada. Para la binomial, el currículo presentó cinco argumentos, excluyendo la validación mediante ejemplo o contraejemplo. En tanto que una serie de libros contempló tres, exceptuando la argumentación a través de generalización y simulaciones con herramienta tecnológica y objeto manipulable, y la otra serie sólo presentó dos, descartando aquellas tres y la justificación mediante gráfica.

En cuanto que en el contexto de la normal los lineamientos curriculares promovieron cuatro tipos de argumentos, omitiendo la justificación mediante ejemplo o contraejemplo, Mientras que una serie de libros sugirió tres, descartando la validación mediante simulación con herramienta tecnológica y generalización, y la otra serie contempló cuatro, suprimiendo la justificación por medio de representación gráfica. De esta manera, se constató que existen objetos matemáticos primarios vinculados a la variable aleatoria y distribuciones binomial y normal que han sido identificados en el currículo chileno, pero que fueron excluidos de los libros de texto, y viceversa, demostrándose una falta de coherencia entre las directrices propuestas por el MINEDUC (2015; 2019a) y aquel recurso.

Aún más, los análisis desarrollados en el Estudio 2 y el Estudio 3 expuestos en el Capítulo 3, han permitido establecer: i) el significado institucional de referencia de variable aleatoria y distribuciones binomial y normal según el currículo escolar chileno; y ii) el significado institucional pretendido de variable aleatoria y distribuciones binomial y normal según los libros de texto analizados.

Por último, los resultados tanto del Estudio 2 como del Estudio 3 confirman la hipótesis que en el tratamiento de variable aleatoria y distribuciones binomial y normal existen diferencias entre el sugerido por la normativa curricular chilena y propuesto por los libros de texto chilenos. Por una parte, se observaron discrepancias en las situaciones-problemas (actividades, tareas, ejercicios y problemas) sobre los temas de interés, sugeridas en los lineamientos curriculares chilenos y las promovidas en los libros de Chile analizados. Por otra parte, se han evidenciado discordancias entre los documentos analizados en torno a: i) las proposiciones, procedimientos y argumentos vinculados a la variable aleatoria; ii) el lenguaje, conceptos, procedimientos y argumentos relacionados con la distribución binomial; y iii) los conceptos, proposiciones, procedimientos y argumentos vinculados a la distribución normal.

6.2.3 Conclusiones sobre el tercer Objetivo e Hipótesis

(O3): *Construir y validar un instrumento orientado a evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en el contexto escolar chileno*.

Este objetivo ha sido alcanzado a través de la realización de dos estudios (Estudio 4 y Estudio 5) presentados en el Capítulo 4. El primero de ellos, ha abordado la construcción del instrumento y su validez de contenido, para lo cual se siguió un proceso metodológico riguroso que incluyó diversas fases, mientras que el segundo, presentó la validez de constructo y fiabilidad del instrumento.

En el Estudio 4 (Bizet et al., 2023c), primero fue creada la Guía de Situaciones-Problemas sobre Variable Aleatoria y sus Aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar Chileno (GSP-VADP), cuyo proceso de construcción se organizó en dos subetapas. La primera subetapa abordó la identificación de campos de problemas sobre los temas de interés a partir de un análisis epistemológico. En concreto, con base en el análisis de investigaciones sobre aspectos epistemológicos de la variable aleatoria, distribución binomial y distribución normal, expuesto en el Capítulo 2, y mediante un proceso cíclico e inductivo, se establecieron ocho campos de problemas ($C\text{-}P_1$ a $C\text{-}P_8$) ligados a los temas en cuestión.

La segunda subetapa para construir la GSP-VADP abordó el reconocimiento de situaciones-problemas sobre los temas de interés desde un análisis de lineamientos curriculares. Puntualmente, a partir de un análisis de contenido al currículo escolar chileno, los lineamientos curriculares americanos y el conocido proyecto GAISE, se identificaron normativas vigentes (unidades de análisis) en torno a la enseñanza-aprendizaje de la variable aleatoria y distribuciones de probabilidad. Luego las unidades de análisis fueron clasificadas en los ocho campos de problemas (categorías de análisis) previamente definidos. Finalmente, aquellas unidades de análisis han sido comparadas y reducidas, y a partir de estas se infieren 33 situaciones-problemas sobre variable aleatoria, distribución binomial y distribución normal, que en su conjunto conforman la GSP-VADP.

En la segunda fase del Estudio 4, a partir de la literatura se eligió un conjunto inicial de ítems que componen el instrumento de evaluación. Específicamente para cada situación-problema de la GSP-VADP se propusieron tres ítems, que fueron seleccionados de investigaciones previas, tareas propuestas en libros de texto previamente analizados (Capítulo 3) o elaborados por los autores. Cabe destacar que en esta fase se excluyeron dos situaciones-problemas (S-P$_{6.5}$ determinar probabilidades asociadas a una distribución binomial empleando una herramienta tecnológica y S-P$_{7.5}$ determinar probabilidades asociadas a una distribución normal empleando una herramienta tecnológica), debido a que se proyectaba que el instrumento fuese aplicado a estudiantes individualmente y en forma presencial o en línea.

En la tercera fase, fue elegido el conjunto final de ítems que componen el instrumento, mediante la validez de contenido por juicio de expertos. Participaron seis doctores en educación matemática con experiencia en investigación en didáctica de la estadística, quienes evaluaron de forma independiente tres ítems para cada tarea. Así se obtuvo la versión inicial del instrumento, seleccionándose en cada situación-problema el ítem con media más alta y menor desviación estándar. Además, a cada uno de estos se le calculó su Coeficiente de Validez de Contenido (Hernández-Nieto, 2002) para analizar el grado en que se ajustó a una situación-problema, obteniéndose que el instrumento posee un coeficiente de validez y concordancia bueno (0,87).

En el Estudio 5 (Bizet et al., en evaluación), la versión inicial del instrumento (32 ítems) se aplicó en una muestra piloto de 80 estudiantes universitarios chilenos, quienes no habían cursado en la universidad alguna asignatura sobre estadística y probabilidad. Luego fue indagada la validez de constructo del instrumento mediante un análisis factorial exploratorio (AFE) y confirmatorio (AFC), y su fiabilidad empleando el coeficiente de alfa de Cronbach.

Los resultados evidenciaron en el AFE, que la estructura empírica del cuestionario está compuesta por seis factores que explican el 58% de la varianza total, cada uno relacionado con un campo de problema (constituido por tres a cuatro ítems) sobre los temas en cuestión: i) el primer factor se podría asociar al C-P$_4$ Establecer la función de distribución de una variable aleatoria discreta; ii) el segundo factor al C-P$_2$ Reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria; iii) el tercer factor al C-P$_6$ La distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real; iv) el cuarto factor al C-P$_5$ Utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria; v) el quinto factor al C-P$_7$ La distribución normal como un modelo probabilístico que describe un fenómeno de la vida real; y vi) el sexto factor posible de vincularse al C-P_1 Identificar la variable aleatoria como una función presente en el contexto probabilístico.

De esta manera, en la estructura empírica del instrumento fueron descartados dos campos de problemas: i) el C-P_3 vinculado a la función de probabilidad de una variable aleatoria continua, aunque parte de su contenido se incluyó en el C-P_1, mediante el $I_{3.1}$ representante de la S-$P_{3.1}$ sobre graficar aquella función; y ii) el C-P_8 relacionado con la aproximación de distribuciones de variables aleatorias discretas, aunque se abordó la aproximación de la binomial por la normal en el C-P_7, mediante el $I_{7.3}$ representante de la S-$P_{7.3}$ sobre describir la tendencia de datos representados gráficamente empleando aquella aproximación.

En el AFC, los resultados evidenciaron sobre los coeficientes factor-ítem que la mayoría poseen valores adecuados. Respecto a los coeficientes factor-factor, se obtuvo que existe relación media entre las respuestas al C-P_2 vinculado a la función de probabilidad de una variable aleatoria discreta, con las respuestas al C-P_1 que involucra la variable aleatoria como función, y con las respuestas al C-P_4 en torno a la función de distribución de una variable aleatoria discreta. Además hay relación media entre los campos de problemas sobre modelos probabilísticos, es decir, entre el C-P_6 distribución binomial y el C-P_7 distribución normal. Mientras que existe relación baja entre las respuestas: i) al C-P_1 vinculado a la variable aleatoria como función y C-P_5 relativo a algunos valores de posición central o de dispersión asociados a la variable aleatoria; y ii) al C-P_2 sobre la función de probabilidad de una variable aleatoria discreta y C-P_6 relacionado con la distribución binomial.

También en el AFC se obtuvo que la estructura empírica del instrumento constituida por seis factores y 19 ítems tiene medidas de bondad de ajuste aceptables. Por otra parte, fue calculada el alfa de Cronbach del instrumento, cuyo valor obtenido demostró que el cuestionario es fiable y existe consistencia entre los ítems y constructo.

El Estudio 3 y Estudio 4 se complementaron con el cálculo del índice de dificultad y discriminación de cada uno de los 19 ítems que componen la versión final del instrumento. Los resultados mostraron que la mayoría de los ítems poseen una adecuada discriminación, además tres ítems presentaron una dificultad moderada y 16 fueron difíciles. Por tanto, los resultados conseguidos en el Capítulo 4 garantizan que la versión final del cuestionario (19 ítems) es válida (posee validez de contenido y constructo) y fiable para indagar la compresión de variable aleatoria y distribuciones binomial y normal en egresados de educación escolar chilena. Además, el cuestionario diseñado caracteriza tanto el significado institucional evaluado de variable aleatoria como el significado institucional evaluado de distribuciones binomial y normal.

De esta forma, los resultados obtenidos en el Capítulo 4 confirman la hipótesis que el instrumento diseñado para evaluar la comprensión de variable aleatoria y distribuciones binomial y normal posee buenas características psicométricas. En este sentido se destaca la adecuada consistencia interna del cuestionario, al igual que su estructura factorial de seis factores.

6.2.4 Conclusiones sobre el cuarto Objetivo e Hipótesis

(O4): *Evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en estudiantes egresados de educación escolar chilena.*

Este objetivo fue logrado a través del Estudio 6 (ver Capítulo 5), donde se evalúo la comprensión de variable aleatoria y modelos de probabilidad en 101 estudiantes egresados de educación escolar chilena, mediante la aplicación de un cuestionario previamente validado (Capítulo 4) e integrado por 19 ítems. Los resultados globales de este estudio mostraron que el puntaje total obtenido por cada participante varió entre 0 y 20 puntos (sobre una puntuación máxima de 38 puntos) y la puntuación media alcanzada fue 4 puntos.

Sobre la comprensión de variable aleatoria, los resultados evidenciaron que la mayoría de las situaciones-problemas valoradas en el cuestionario han sido resueltas por los participantes, pero en un bajo porcentaje. Las tareas en las que se evidenciaron mayor proporción de respuestas parciales o correctas han sido: i) la S-$P_{1.1}$ diferenciar entre variables aleatorias y variables con dependencia funcional (43%) y la S-$P_{3.1}$ representar en lenguaje gráfico la función de densidad de una variable aleatoria continua (17,82%), pertenecientes al C-P_1; y ii) la S-$P_{2.2}$ determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde el enfoque de probabilidad clásico (17,82%,) que integra el C-P_2.

En tanto que las tareas donde se obtuvieron menor porcentaje de respuestas parcialmente correctas o correctas fueron: i) la S-$P_{4.3}$ representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta (1%) que integra el C-P_4; y ii) la S-$P_{5.4}$ identificar la media y desviación estándar de una variable aleatoria continua (1%) y la S-$P_{5.5}$ interpretar la desviación estándar de una variable aleatoria discreta (0%), pertenecientes al C-P_5.

Además en las tareas sobre valorar la comprensión de variable alearia fueron observados ciertos conflictos semióticos cognitivos, principalmente en las respuestas parciales y errónea, tales como: confundir el recorrido con el dominio de una variable aleatoria, asumir equiprobabilidad de sucesos aleatorios compuestos, confundir la función de probabilidad con la función de densidad (o viceversa), confundir la media con la moda de una variable aleatoria discreta; argumentar una respuesta a partir de creencias personales sobre el contexto de la tarea, entre otros.

En sintesis el significado personal de variable aleatoria logrado por los participantes carece de ciertos objetos matemáticos primarios que lo diferencian del significado institucional evaluado. Por ejemplo, la situación-problema relativa a interpretar la desviación estándar de una variable aleatoria discreta, ninguno de los jóvenes logró responderla. Aún más limitados estudiantes de la muestra han logrado emplear los distintos objetos matemáticos involucrados en la solución de las tareas evaluadas. En consecuencia los estudiantes chilenos egresados de educación escolar posee una comprensión baja de variable aleatoria.

En cuanto a la comprensión de los modelos de probabilidad, los resultados mostraron que todas las situaciones-problemas evaluadas en el cuestionario han sido resueltas por los participantes, aunque en una baja proporción. Las tareas en las que se evidenciaron mayor porcentaje de respuestas parciales o correctas han sido: i) la S-$P_{6.4}$ determinar probabilidades asociadas a una distribución binomial de forma manual (15%) que integra el C-P_6; y ii) la S-$P_{7.4}$ evaluar la veracidad de afirmaciones utilizando el concepto de distribución binomial (43,56%.) perteneciente al C-P_7. Mientras que las tareas donde se lograron menor porcentaje de respuestas parcialmente correctas o correctas fueron: i) la S-$P_{6.1}$ identificar situaciones que pueden modelizarse mediante una distribución binomial (1%) y la S-$P_{6.3}$ calcular la media de una distribución binomial (1%) pertenecientes al C-P_6; y ii) la S-$P_{7.2}$ determinar probabilidades asociadas a una distribución normal de forma manual (1%) que integra la S-$P_{7.}$

Asimismo, en las tareas en torno a valorar la comprensión de los modelos de probabilidad fueron observados algunos conflictos semióticos cognitivos, principalmente en las respuestas parciales y errónea, por ejemplo: asumir equiprobabilidad de sucesos aleatorios compuestos, confundir la media con la mediana de una variable aleatoria discreta y traducir del lenguaje verbal al simbólico incorrectamente en el cálculo de probabilidades, entre otros.

En resumen, el significado personal de distribución binomial logrado por los estudiantes de la muestra es coherente con el significado institucional evaluado en el cuestionario, aunque pocos participantes lograron utilizar la diversidad de objetos matemáticos primarios asociados a la resolución de las situaciones-problemas valoradas. Por tanto, los participantes poseen una comprensión baja de distribución binomial.

Finalmente, el significado personal de distribución normal logrado por los jóvenes participantes carece de objetos matemáticos primarios que lo diferencian del significado institucional evaluado, tales como: el lenguaje tabular; los conceptos de media, desviación estándar, simetría y distribución normal estándar; la propiedad de los intervalos centrales, aquellas para calcular probabilidades con distribución normal estándar y las condiciones para aproximar una binomial por una normal; y el proceso de calcular probabilidades, la tipificación y la lectura de la tabla de la normal estándar. También pocos estudiantes de la muestra lograron aplicar algunos objetos primarios asociados a la solución de las tareas valoradas. Por consiguiente, los participantes poseen una comprensión baja de la distribución normal.

De esta forma, los resultados obtenidos en el Capítulo 5 confirmar la hipótesis que los participantes tienen una comprensión insuficiente de variable aleatoria y distribuciones binomial y normal. En este sentido se destacan los resultados evidenciados en cada campo de problema. En el C-P_1 identificar la variable aleatoria como una función presente en el contexto probabilístico, un participante logró la puntuación máxima (6 puntos) aunque la puntuación media lograda fue de 1,2 puntos. Para el C-P_2 reconocer la distribución de probabilidad de una variable aleatoria discreta como una herramienta que permite ver la variación aleatoria, ningún estudiante alcanzó la puntuación máxima (8 puntos) y la puntuación media obtenida ha sido de 1,2 puntos. También en el C-P_4 establecer la función de distribución de una variable aleatoria discreta y el C-P_5 utilizar algunos valores de posición central o de dispersión vinculados a la variable aleatoria, ninguno de los participantes consiguió la puntuación máxima (6 puntos) y la puntuación media lograda fue de 0,1 y 0,2 puntos respectivamente

Además, en el C-P_6 la distribución binomial como un modelo probabilístico que describe un fenómeno de la vida real, un participante consiguió la puntuación máxima (6 puntos), pero la puntuación media lograda fue de 0,7 puntos. Por último, para el C-P_7 la distribución normal como un modelo probabilístico que describe un fenómeno de la vida real, ninguno de los estudiantes alcanzó la puntuación máxima (6 puntos) y la puntuación media obtenida fue de 0,5 puntos

6.3 Principales aportaciones

El análisis de los lineamientos curriculares principalmente de Chile y Estados Unidos ha permitido identificar las principales tareas sobre variable aleatoria y distribuciones binomial y normal que se sugieren a nivel escolar. Así proponer una Guía de Situaciones-Problemas sobre Variable Aleatoria y sus Aplicaciones en Distribuciones de Probabilidad según el Currículo Escolar (Bizet et al., 2023c), la cual posee información valiosa que puede servir a los profesores de matemáticas que ejercen su labor en las escuelas, para orientar su trabajo y diseñar propuestas de enseñanzas en torno a los temas en cuestión.

Los estudios de los libros de texto chilenos han permitido ampliar los tipos de actividades sobre variable aleatoria y modelo normal presentes en investigaciones previas en torno a este recurso didáctico (Ortiz, 2002; Valverde, 2017). Además, hicieron posible complementar los conceptos, lenguajes, procedimientos, proposiciones y argumentos relacionados con la variable aleatoria y la distribución normal reconocidos aisladamente en estudios previos de textos escolares (Doukhan y Guedet, 2019; Ortiz, 2002, Valverde, 2017). Otro aporte novedoso, fue la identificación de las tareas sobre el modelo binomial y diversos objetos matemáticos primarios que intervienen en su solución, que promueven los textos escolares, pues no se han encontrado estudios que aborden el tratamiento de la binomial en este recurso didáctico, diferenciando los diversos objetos matemáticos involucrados.

De esta manera, los estudios sobre libros de texto (Bizet et al., 2023a; 2023b) pueden ser un insumo valioso para los profesores, que deben enseñar en escuelas los tópicos de variable aleatoria y/o distribuciones binomial y normal. También aquellos estudios pueden tener información valiosa para los autores de textos escolares, ya que aquellos estudios proporcionan criterios para mejorar la elaboración de libros de texto, los cuales continúa siendo el recurso didáctico más empleado para enseñar matemáticas en la educación escolar (Vásquez y Alsina, 2015).

Por otra parte, se ha aportado con un cuestionario validado para evaluar la comprensión de variable aleatoria y distribuciones binomial y normal en estudiantes egresados de educación escolar (Bizet et al., en evaluación), fundamentado en lineamientos curriculares y el Enfoque

Ontosemiótico del Conocimiento y la Instrucción Matemáticos. Además, este instrumento podría tener diversas utilidades: i) evaluar la comprensión de estudiantes escolares posterior a un proceso de enseñanza-aprendizaje de variable aleatoria y/o modelos de probabilidad; ii) diagnosticar la comprensión de estudiantes que acaban de ingresar a la universidad sobre dichos temas; o iii) valorar el conocimiento matemático de variable aleatoria y modelos binomial y normal en futuros profesores de matemáticas ya que son temas que deberán enseñarlos en la escuela.

En cuanto al estudio de evaluación de la comprensión, ha aportado con establecer el nivel de comprensión de variable aleatoria y distribuciones binomial y normal que poseen estudiantes egresados de educación escolar chilena, con base en el EOS. También dicho estudio ha contribuido a establecer el significado personal de los temas de interés logrado por los participantes y los conflictos semióticos que ellos manifiestan al trabajar con la variable aleatoria y los modelos de probabilidad. Esta información en su mayoría es original entonces podría aportan al desarrollo de la educación estadística.

6.4 Limitaciones y futuras líneas de investigación

Respecto a las limitaciones de esta investigación se han identificado algunas en sus diferentes fases. En los estudios sobre análisis de libros de texto chilenos, una restricción ha sido la cantidad de series utilizadas, por lo que si se amplía aquellas series podría conseguirse una mejor caracterización sobre el tratamiento de variable aleatoria y distribuciones binomial y normal en este importante recurso. Otra limitación del trabajo fue omitir la indagación del recurso *cuaderno de actividades* que posee cada libro de texto, pues no se disponían de todos.

En el estudio sobre validez de constructo del instrumento, una de las limitaciones ha sido el tamaño de la muestra, debido a la emergencia sanitaria originada por la pandemia de la Covid-19 fue muy difícil acceder a estudiantes chilenos de primer año universitario. Otro de los inconvenientes de este estudio fue la forma de implementar el instrumento, pues producto de la Covid-19 se debió aplicar el cuestionario en línea durante una sesión de 90 minutos. En este sentido serían necesario un estudio complementario donde se utilice una muestra de mayor tamaño y representatividad, e implemente el cuestionario de manera presencial, para afianzar o refutar los resultados obtenidos en el Estudio 5.

En el estudio sobre evaluación de la comprensión, una de las restricciones fue el tipo de muestra (no probabilística). Esta ha sido elegida de manera intencionada según la disponibilidad de universidades chilenas que participaron en el estudio, donde producto de la pandemia de la Covid-19 fue complejo que los profesores cedieran una sesión en línea de su asignatura, y los estudiantes accedieran a responder en línea a un cuestionario de preguntas abiertas y enviar fotografías de su trabajo. Otra limitación fue la manera de implementar el instrumento, pues producto que el cuestionario fue aplicado en línea, la mayoría de los participantes que respondieron a las tareas omitieron su justificación, aunque fue solicitada.

Con relación a las futuras líneas de investigación, a continuación se proyectan algunas. Puntualmente en los libros de texto sería deseable identificar los potenciales conflictos semióticos en torno al tratamiento de la variable aleatoria y distribuciones binomial y normal.

Por otra parte, los resultados del estudio de evaluación plantean el desafío de realizar una investigación que aborde el diseño e implementación de una unidad didáctica, dirigida a estudiantes escolares de 16 a 18 años, que introduzca paulatinamente los campos de problema sobre variable aleatoria y modelos de probabilidad. Posteriormente analizar si se genera una mejora en la comprensión de los temas, mediante la implementación del cuestionario diseñado.

Finalmente, sería interesante investigar la comprensión de variable aleatoria y distribuciones binomial y normal en futuros profesores de matemáticas, incluyendo tanto el conocimiento matemático como didáctico de los temas. Debido a que estos son propuestos en lineamientos curriculares y en un futuro próximo deben enseñarlos en las escuelas.

Referencias

Alvarado, H. (2007). *Significado del teorema central del límite en la enseñanza de la estadística en ingeniería* [Tesis de Doctorado, Universidad de Granada].

Alvarado, H. y Batanero, C. (2007). Dificultades de Comprensión de la Aproximación Normal a la Distribución Binomial. *Números, revista de didáctica de las matemáticas*, 67, 1-7.

Alvarado, H. y Batanero, C. (2008). El significado del teorema central del límite en textos universitarios de probabilidad y estadística. *Estudios Pedagógicos, 34*(2),7-28.

Alvarado, H., Estrella, S., Retamal, L. y Galindo, M. (2018). Intuiciones probabilísticas en estudiantes de ingeniería: implicaciones para la enseñanza de la probabilidad. *Revista Latinoamericana de Investigación en Matemática Educativa*, *21*(2), 131-156.

Alvarado, H. y Retamal, L. (2010). La aproximación binomial por la normal: Una Experiencia de Reflexión Sobre la Práctica. *Revista Paradigma*, *31*(2), 89-108.

Alvarado, H. y Retamal, M. (2014). Representaciones de la distribución de probabilidad binomial. En L. Andrade (Ed.), *Memorias del I Encuentro Colombiano de Educación Estocástica* (pp. 98-109). Asociación Colombiana de Educación Estocástica.

Amrani, H. y Zaki, M. (2015). Student's conceptual difficulties with respect to the notion of random variable. *International Journal of Education, Learning and Development, 3*(9), 82-98.

Armah, G., Asiedu-Addo, S. y Owusu-Ansah, N. (2016). An investigation on the conceptual understanding of the standard deviation of fresh undergraduate students. *ASIO Journal of Chemistry, Physics, Mathematics and Applied Sciences*, *1*(3), 1-9.

Ball, D., Thames, M. y Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59(*5), 389-407.

Bansilal, S. (2012). Using conversions and treatments to understand students' engagement with problems based on the normal distribution curve. *Pythagoras, 33*(1), 1-13.

Bansilal, S. (2014). Using an apos framework to understand teachers' responses to questions on the normal distribution. *Statistics Education Research Journal*, *13*(2), 42-57.

Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L. y Spangler, D. (2020). *Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II* (GAISE II). ASA.

Barraza, M. (2007). La consulta a expertos como estrategia para la recolección de evidencias de validez basado en el contenido. *Investigación Educativa Duranguense, 2*(7), 5-14.

Batanero, C. (2004). Ideas Estocásticas fundamentales. ¿Qué contenido se debe enseñar en la clase de probabilidad? En J. Fernandes, M. Sousa y S. Ribeiro (Eds.), *Ensino e aprendizagem de probabilidades e estatística – Actas do I Encontro de Probabilidades e Estatística na Escola* (pp. 9-30). Centro de Investigação em Educação da Universidade do Minho.

Batanero, C., Chernof, E., Engel, J., Lee, H. y Sánchez, H. (2016). *Research on Teaching and Learning Probability*. Springer.

Batanero, C. y Díaz, C. (2005). Análisis del Proceso de Construcción de un Cuestionario sobre Probabilidad Condicional. Reflexiones desde el Marco de la TFS. En L. Ordóñez, C. Batanero y Á. Contreras (Eds.), *Investigación en didáctica de la matemática*

Batanero, C., Gea, M., Díaz- Levicoy, D. y Cañadas, G. (2015). Objetos matemáticos ligados a la regresión en los textos españoles de bachillerato. *Educación Matemática*, *27*(2), 9-35.

Batanero, C., Tauber, L. y Sánchez, V. (2001). Significado y comprensión de la distribución normal en un curso introductorio de análisis de datos. *Quadrante, 10*(1), 59-92.

Batanero, C., Tauber, L. y Sánchez, V. (2004). Students' reasoning about the normal distribution. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 257–276). Springer.

Begué, N., Batanero, C., Gea, M. y Díaz-Levicoy, D. (2020). Razonamiento de estudiantes de bachillerato ante una situación binomial. *Tangram*, *3*(2), 27-50.

Ben-Zvi, D. y Garfield, J. (2004). Statistical literacy, reasoning and thinking: Goals, definitions andchallenges. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3- 16). Springer.

Berrío, J., González, A. Meriño, V. y Valbuena, S. (2022). ¿Cómo interpretan noticias sobre el covid-19 los estudiantes de un programa de formación de profesores de matemáticas? En J. Solorzano y D. Suárez (Eds.), *Miradas y perspectivas de la educación matemática desde la formación, la inclusión y la tecnología* (pp. 146-173). Sello Editorial Coruniamericana.

Biggs, J. y Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. Academic Press.

Bill, A., Watson, K. y Gayton, P. (2009). Guessing Answers to Pass a 5- item True False Test: Solving a Binomial Problem Three Different Ways. En R. Hunter, B. Bicknell y T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conferencia del Grupo de Investigación en Educación en Matemáticas de Australasia* (pp. 57-64). MERGA.

Bizet, V., Molina-Portillo, E., Ruz, F. y Contreras, J.M. (2022). What does the research tell us about the understanding of the random variables and its probability distribution? *Boletín de Estadística e Investigación Operativa*, *38*(3), 1-22.

Bizet, V., Molina-Portillo, E. y Contreras, J.M. (2023a). Situaciones-problemas sobre variable aleatoria y sus aplicaciones en distribuciones de probabilidad según libros de texto chilenos. *Profesorado, revista de currículum y formación de profesorado, 27*(2), 351-382.

Bizet, V., Molina-Portillo, E. y Contreras, J.M. (2023b). Objetos matemáticos ligados a la variable aleatoria y sus distribuciones de probabilidad en libros de texto chilenos. *PNA, revista de investigación en Didáctica de la Matemática*, *17*(2), *201-238.*

Bizet, V., Molina-Portillo, E., Ruz, F. y Contreras, J.M. (2023c). Elaboración de una guía de situaciones-problema sobre variable aleatoria y sus aplicaciones a partir del currículo escolar chileno. *Educación Matemática, 35*(1), 163-190.

Bizet, V., Molina-Portillo, E. y Contreras, J.M. (en evaluación). Cuestionario para valorar la comprensión de variable aleatoria y distribuciones de probabilidad en egresados de educación escolar.

Bizet, V. y Ramos, E. (2019). Una experiencia de enseñanza para abordar la variable aleatoria con estudiantes de secundaria. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística* (pp. 1-10). Departamento de didáctica de la matemática de la Universidad de Granada.

Bizet, V. y Ramos, E. (2022). Valoración de una situación didáctica para la enseñanza de variable aleatoria y distribución de probabilidad en la educación secundaria chilena. *Innovaciones Educativas*, *24*(36), 21–36.

Blanco, M., Bozt, J., Calderón, F., Jiménez, M., González, M., López, G., Romero, P., Díaz, M., Muñoz, G. y Rupin, P. (2009). *Matemática 2 proyecto bicentenario*. Santillana.

Blanco, M., Bozt, J., Calderón, F., Romero, L., Jiménez, L. y Jammet, C. (2009). *Matemática 3 proyecto bicentenario.* Santillana.

Brousseau, G. (1997). *Theory of didactical situations in mathematics. Didactique des mathématiques, 1970-1990*. Kluwer.

Brousseau, G. (2007). *Iniciación al estudio de la teoría de las situaciones didácticas*. Libros del Zorzal.

Burrill, G. y Biehler, R. (2011). Fundamental Statistical Ideas in the School Curriculum and in Training Teachers. En C. Batanero, G. Burrill, y C. Reading (Eds.), *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study* (pp.57-69). Springer.

Calandra, M. y Costa, V. (2015). La problemática de la enseñanza y aprendizaje del concepto de variable aleatoria continua y de función de densidad de probabilidad. En FaHCE (Ed.), *Actas IV jornadas de enseñanza e investigación educativa en el campo de las ciencias exactas y naturales* (pp. 1–10). UNLP.

Calderón, D., García-García, J., Fernández, N. y Hernández E. (2022). Cálculo de probabilidades en tablas de contingencia por estudiantes chilenos de primer año medio. REXE, *21*(47), 50-74.

Carpio, M., Gaita, C., Wilhelmi, M. y Sáenz de Cabezón, A. (2009). Significados de la distribución normal en la universidad. En M. González, M. González y J. Murillo (Eds.), *Investigación en Educación Matemática XIII* (pp. 1-15). SEIEM.

Chacón, A., García, G., Rupín, P., Setz, J. y Villena, M. (2018). *Texto del estudiante de matemática 2° medio*. Ediciones SM.

Chevallard, Y. (1980). The didactics of mathematics: its problematic and related research. *Recherches en Didactiques des Mathématiques, 1*, 146-157.

Chevallard, Y. (1991). *La transposición didáctica: del saber sabio al saber enseñado*: Aique.

Chevallard, Y. (1999). El análisis de las prácticas docentes en la teoría antropológica de los didáctico. *Recherches en Didáctique des Mathématiques*, *19*(2), 221-266.

Choo-Kim, T. y Choo-Peng, T. (2015). Effects of the handheld technology instructional approach on performances of students of different achievement levels. *Computers and Education*, *82*, 306-314.

Cobo, B. (2003). *Significado de las medidas de posición central para los estudiantes de secundaria* [Tesis de Doctorado, Universidad de Granada].

Colomé, S. (28 de diciembre de 2015). *¿Cuál es la probabilidad real del empate de la CUP?.* La Vanguardia. https://www.lavanguardia.com/politica/20151228/301071423295/probabilidad-empate-cup.html

Crews, T., Biswas, G., Goldman, S. y Bransford, J. (1997). Anchored interactive learning environments. *International Journal of Artificial Intelligence in Education*, *8*, 142-178.

De Bock, D., Verschaffel, L. y Janssens, D. (2002). The effects of different problem presentations and formulations on the illusion of linearity in secondary school students. *Mathematical Thinking and Learning*, *4*(1), 65–89.

Del Pino, J. (2019). *Las medidas de dispersión en la educación secundaria obligatoria: análisis de libros de texto y de la comprensión de los estudiantes* [Tesis de Doctorado, Universidad de Jaén].

Del Pino, G. y Estrella, S. (2012). Educación estadística: relaciones con la matemática. *Pensamiento Educativo Revista de Investigación Educacional Latinoamericana*, *49*(1), 53-64.

Departamento de Investigaciones Educativas. (2014). *Matemática 4 proyecto bicentenario*. Santillana.

Derouet, C. (2019). Introduire la notion de fonction de densité de probabilité: dynamiques entre trois domaines mathématiques. *Recherches en Didactique des Mathematiques*, *39*(2), 213-266.

Derouet, C., Planchon,G., Hausberger,T. y Hochmuth, R. (2018). Bridging probability and calculus: the case of continuous distributions and integrals at the secondary-tertiary transition. En V. Durand-Guerrier, R. Hochmuth, S. Goodchild y N. Hogstad (Eds.), *Proceedings of the Second Conference of the International Network for Didactic Research in University Mathematics (pp. 497-506).* University of Agder and INDRUM.

Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología* [Tesis de Doctorado, Universidad de Granada].

Díaz, D. (2018). *Comprensión de gráficos estadísticos por alumnos chilenos de educación primaria* [Tesis de Doctorado, Universidad de Granada].

Díaz, E., Ortiz, N., Morales, K., Rebolledo, M., Barrera, R. y Norambuena, P. (2020). *Texto del estudiante de matemática 2° medio*. Ediciones SM.

Dinges, H. (2005). Variables, in particular random variables. En M. Hoffmann, J. Lenhard, y F. Seeger (Eds.), *Activity and sign grounding mathematics education* (pp. 305–311). Springer.

Doukhan, C. y Gueudet, G. (2019). Students' difficulties at the secondary-tertiary transition: the case of random variables. En U. Jankvist, M. van den Heuvel-Panhuizen y M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 2464-2471). Freudenthal Group y Freudenthal Institute of the Utrecht University y ERME.

Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. En D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 95-126). Kluwer.

Duval, R. (2006). A cognitive analysis of problems of comprehension in the learning of mathematics. *Educational Studies in Mathematics*, *61*, 103–131.

Ebel R. y Frisbie, D. (1991). *Essentials of Educational Measurement* (5ª ed.). Prentice-Hall.

Escobedo, M., Hernández, J., Estebané, V. y Martínez G. (2016). Modelos de ecuaciones estructurales: características, fases, construcción, aplicación y resultados. *Ciencia y Trabajo, 18*(55), 16-22.

Escolano, A. (2009). El manual escolar y la cultura profesional de los docentes. *Tendencias pedagógicas, 14,* 169-180.

Everitt, B. y Wykes, T. (2001). *Diccionario de estadística para psicólogos.* Ariel.

Fernández, F., Andrade, L. y Sarmiento, B. (2013). Rehaciendo el camino hacia la comprensión de la variable aleatoria. En P. Perry, C. Samper, Ó. Molina, L. Camargo, A. Echeverry, F. Fernández, B. Sarmiento (Eds.), *Aportes investigativos para el diseño curricular en geometría y estadística* (pp. 93–169). Universidad Pedagógica Nacional.

Fernández, N., García-García, J., Arredondo, E. e Imilpán, I. (2022). Knowledge of binomial distribution in pre-service mathematics teachers. En S. Peters y L. Zapata-Cardona (Eds.), *11th International Conference on Teaching Statistics* (pp. 1-6). IASE

Flores, B., García, J. y Sánchez, E. (2014). Avances en la calidad de las respuestas a preguntas de probabilidad después de una actividad de aprendizaje con tecnología. En M. González, M. Codes, D. Arnau y T. Ortega (Eds.), *Investigación en Educación Matemática XVIII* (pp. 307-316). SEIEM.

Font, V. y Godino, J. (2006). La noción de configuración epistémica como herramienta de análisis de textos matemáticos: su uso en la formación de profesores. *Educação Matemática Pesquisa, 8*(1), 67-98.

Font, V., Godino, J. y Gallardo, J. (2013). The emergence of objects from mathematical practices. *Educational Studies in Mathematics, 82*(1), 97-124.

Fox, D. (1981). *El proceso de investigación en la educación*. Eunsa.

Franklin, C., Kader, G., Newborn, D., Moreno, J., Peck, R., Perry, M. y Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: a pre-k–12 curriculum framework.* American Statistical Association.

Gal, I. (2002). Adult's statistical literacy: meanings, components, responsibilities. *International Statistical Review, 70*(1), 1– 25.

Gal, I. (2005). Towards "probability literacy" for all citizens: building blocks and instructional dilemmas. En G. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 43-70). Springer.

Gal, I. (2012). Developing probability literacy: needs and pressures stemming from frameworks of adult competencies and mathematics curricula. En S.J. Cho (Ed.), *Proceedings of the 12th International Congress on Mathematical Education* (pp. 1-7). Springer Open.

Galicia, S., Nájera, A. y Sánchez, E. (2013). Niveles de razonamiento frente a problemas binomiales. En J. M. Contreras, G. Cañadas, M. Gea y P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria* (pp. 409-416). Departamento de didáctica de la matemática de la Universidad de Granada.

Galindo-Domínguez, H. (2020). *Estadística para no estadísticos: una guía básica sobre la metodología cuantitativa de trabajos académicos.* 3Ciencias Editorial.

García-García, J., Arredondo, E. y Márquez, M. (2018). Desarrollo de la noción de distribución binomial en estudiantes de bachillerato con apoyo de tecnología. *Revista Paradigma*, *39*(2), 92-106.

García-García, J., Fernández, N., Arredondo, E. e Imilpán, I. (2022). The Binomial Distribution: Historical Origin and Evolution of Its Problem Situations. *Mathematics*, *10*, 1-28.

García-García, J. y Hernández, E. (2018). Niveles de razonamiento probabilístico de estudiantes de bachillerato sobre la noción de la distribución binomial. En L. Serna y D. Páges (Eds.), *Acta Latinoamericana de Matemática Educativa* (pp. 962-969). CLAME.

García, J., Medina, M. y Sánchez, E. (2014). Niveles de razonamiento de estudiantes de secundaria y bachillerato en una situación-problema de probabilidad. *Avances de Investigación en Educación Matemática*, *6*, 5–23.

García, J. y Sánchez, E. (2013). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a una situación básica de variable aleatoria y distribución. En J. M. Contreras, G. R. Cañadas, M. M. Gea, y P. Arteaga (Eds.), *Actas de las jornadas virtuales de didáctica de la estadística, probabilidad y combinatoria* (pp. 417–424). Departamento de didáctica de la matemática de la Universidad de Granada.

Garfield, J., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C. y Zieffler, A. (2008). Learning to reason about statistical models and modeling. En J. Garfield y D. Ben-Zvi, *Developing students' statistical reasoning, connecting research and teaching practice* (pp. 143-163). Springer.

Gea, M. (2014). *La correlación y regresión en bachillerato: análisis de libros de texto y del conocimiento de los futuros profesores* [Tesis de Doctorado, Universidad de Granada].

George, D. y Mallery, P. (2003). *SPSS for windows step by step: a simple guide and reference. 11.0 update.* Allyn and Bacon.

Godino, J. (1999). Implicaciones metodológicas de un enfoque semiótico- antropológico para la investigación en didáctica de las matemáticas. En T. Ortega (Ed.*), Investigación en Educación Matemática III* (pp. 196-212). SEIEM.

Godino, J. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, *22*(2.3), 237-284.

Godino, J. (2012). Origen y aportaciones de la perspectiva ontosemiótica de investigación en didáctica de la matemática. En A. Estepa, A. Contreras, J. Deulofeu, M. Penalva, F. García y L. Ordóñez (Eds.), Investigación en Educación Matemática XVI (pp. 49 - 68). SEIEM.

Godino, J. (2013). *Diseño y análisis de tareas para el desarrollo del conocimiento didáctico-matemático de profesores*. En J. M. Contreras, G. Cañadas, M. Gea & P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria* (pp. 1-15). Departamento de Didáctica de la Matemática de la Universidad de Granada.

Godino, J. (2014). Diseño y análisis de tareas para el desarrollo del conocimiento didáctico-matemático de profesores. En J. M. Contreras, G. Cañadas, M. Gea & P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y*

Combinatoria (pp. 1-15). Departamento de Didáctica de la Matemática de la Universidad de Granada.

Godino, J. (2017). Construyendo un sistema modular e inclusivo de herramientas teóricas para la educación matemática. En J. Contreras, P. Arteaga, G. Cañadas, M. Gea, B. Giacomone y M. López-Martín (Eds.), *Actas del Segundo Congreso International Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos* (pp. 1-20). Universidad de Granada.

Godino, J. (2018). Bases semióticas, antropológicas y cognitivas del enfoque ontosemiótico en educación matemática. Disponible en http://enfoqueontosemiotico.ugr.es/JDGodino_bases_sac_EOS.pdf.

Godino, J. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques,14*(3), 325-355.

Godino, J., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education, 39*(1-2), 127-135.

Godino, J., Batanero, C. y Font, V. (2008). Un enfoque ontosemiótico del conocimiento y la instrucción matemáticos. *Acta Scientiae, 10*(2), 7-37.

Godino, J., Batanero, C. y Font, V. (2019). The onto-semiotic approach: implications for the prescriptive character of didactics. *For the Learning of Mathematics, 39*(1), 37- 42.

Godino, J., Batanero, C. y Font, V. (2020). El enfoque ontosemiótico: implicaciones sobre el carácter prescriptivo de la didáctica. *Revista chilena de educación matemática, 12*(2), 3-15.

Gómez, E. (2014). *Evaluación y desarrollo del conocimiento matemático para la enseñanza de la probabilidad en futuros profesores de educación primaria* [Tesis de Doctorado, Universidad de Granada].

González, Y. y Ojeda, A. (2017). Comprensión de la distribución normal en bachillerato. En L. Serna (Ed.), *Acta Latinoamericana de Matemática Educativa, 30* (pp. 207-217). CLAME.

González, J., Ojeda, A. y Palacios, J. (2018). Comprensión de profesores de la distribución normal. En D. García e I. Pérez (Eds.), Acta Latinoamericana de Matemática Educativa, 31(2), (pp. 1764-1772). Ciudad de México: CLAME.

Guerrero, H., Batanero, C. y J. M. Contreras. (2016). Conocimientos sobre esperanza matemática en alumnos de bachillerato. En F. España (Ed.), *Actas del XVI congreso de enseñanza y aprendizaje de las matemáticas* (pp. 26-35). Sociedad Andaluza de Educación Matemática THALES.

Guerrero, H., Ortiz, J. y Contreras, J.M. (2017). Evaluación del conocimiento sobre esperanza matemática y juegos equitativos en estudiantes de bachillerato. *Avance de Investigación en Educación Matemática, 11*, 107-125.

Gueudet, G. y Trouche, L. (2009). Towards new documentation systems for mathematics teachers? *Educational Studies in Mathematics*, *71*, 199–218.

Hausberger, T., Derouet, C., Hochmuth, R. y Planchon, G. (2021). Compartmentalisation of Mathematical Sectors: the Case of Continuous Probability Distributions and Integrals. *International Journal of Research in Undergraduate Mathematics Education,7*, 490–518.

Heitele, D. (1975). An epistemological view of fundamental stochastic ideas. *Educational Studies in Mathematics*, *6*, 187–205.

Hernández-Nieto, R. (2002). *Contributions to Statistical Analysis.* Universidad de los Andes.

Hernández, R., Fernández, C. y Baptista, L. (2014). Metodología de la investigación (6.a ed.). McGraw Hill.

Hugues, H. (2005). Uso de Hojas Electrónicas en la Enseñanza de la Distribución Norma. En J. Lezama, M. Sánchez y J. Molina (Eds.), *Acta Latinoamericana de Matemática Educativa, 18* (pp. 757–763). CLAME.

Hwang, H. J. y Chae, J. H. (2020). Analysis of the Secondary Pre-service Mathematics Teachers' Mathematical Knowledge for Teaching: Focused on Normal Distribution. *Journal of the Korean School Mathematics Society, 23*(4), 427-448.

Jones, G., Langrall, C. y Mooney, E. (2007). Research in probability: Responding to classroom realities. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909-955). Information Age Publishing y NCTM.

Kachapova, F. (2012). A general approach to teaching random variables. *Mathematics teaching-research journal online, 5*(2), 1-16.

Kachapova, F. y Kachapov, I. (2012). Students' misconceptions about random variables. *International Journal of Mathematical Education in Science and Technology, 41*(3), 963-971.

Kazak, S. y Pratt, D. (2017). Pre-service mathematics teachers' use of probability models in making informal inferences about a chance game. *Statistics Education Research Journal*, *16*(2), 287-304.

Kline, R. (2005). *Principles and practice of structural equation modeling*. The Guilford Press.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, *6*, 59-98.

Krippendorff, K. (1990). *Metodología de análisis de contenido: teoría y práctica*. Paidós.

Kuzniak, A., Tanguay, D. y Elia, I. (2016). Mathematical Working Spaces in schooling: an introduction. *ZDM, Mathematics Education*, *48*(6), 721-737.

Landín, P. (2013). Niveles de razonamiento probabilístico de estudiantes de bachillerato sobre problemas binomiales En J. M. Contreras, G. Cañadas, M. Gea y P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria* (pp. 425-431). Departamento de didáctica de la matemática de la Universidad de Granada.

Landín, P. y Sánchez, E. (2010). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a tareas de distribución binomial. *Educación Matemática Pesquisa, 12*(3), 598-618.

Lavidas, K., Barkatsas, T., Manesis, D. y Gialamas, V. (2020). A structural equation model investigating the impact of tertiary students' attitudes toward statistics, perceived competence at mathematics, and engagement on statistics performance. *Statistics Education Research Journal*, *19(*2), 27–41.

Lecoutre, M. (1985). Effect d' informations de nature combinatoire et de nature fréquentielle sur le judgements probabilistes. *Recherches en Didactique des Mathématiques, 6,* 193- 213.

Lee, H. (2016). Probability Concepts Needed for Teaching a Repeated Sampling Approach to Inference. En C. Batanero y E. Chernoff (Eds.), Teaching and Learning Stochastics, Advances in Probability Education Research (pp. 89-101). Springer.
Lee, H. S., Angotti, R. y Tarr. J. (2010). Making comparisons between observed data and expected outcomes: students' informal hypothesis testing with probability simulation tools. *Statistics Education Research Journal, 9*(1), 68-96.
Li, J., Cheng, J., An, T. y Zhou, D. (2021). Comparative Study on Binomial Distribution Content in High School Textbooks. En J. Wang (Ed.), *School mathematics textbooks in China, comparative studies and beyond* (pp. 107-136). Word scientific.
Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología, 30*(3), 1151-1169.
Lugo-Armenta, J. y Pino-Fan, L. (2021). Niveles de Razonamiento Inferencial para el Estadístico t-Student. *Bolema, 35*(71), 1776-1802.
Marco-Buzunáriz, M. A., Muñoz-Escolano, J. M. y Oller-Marcén, A. M. (2016). Investigación sobre libros de texto en los simposios de la SEIEM (1997-2015). En J. Macías, A. Jiménez, J. González, M. Sánchez, P. Hernández, C. Fernández, F. Ruiz, T. Fernández y A. Berciano (Eds.), *Investigación en Educación Matemática XX* (pp. 325-334). SEIEM.
Martínez, M. (2020). La mejora de los logros escolares: la medida y valoración de la inspección educativa. *Educação e Pesquisa*, *46*, 1-17.
Maurer, K., Hudiburgh, L. y Werwinski, L. (2020). What do students gain from games? Dice gamesvs word problems. *Teaching Statistics, 42*(2), 41-46.
Mayén, S. (2009). *Comprensión de las medidas de tendencia central en estudiantes mexicanos de educación secundaria y bachillerato* [Tesis de Doctorado, Universidad de Granada].
Mevarech, Z. y Kramarsky, B. (1997). From verbal descriptions to graphic representations: stability and change in students' alternative conceptions. *Educational Studies in Mathematics*, *32*(3), 229–263.
Miller, T. K. (1998). The random variable concept in introductory statistics. En L. Pereira-Mendoza (Ed.). *Proceedings of the Fifth International Conference on Teaching Statistics* (pp.1221-1222). IASE.
Ministerio de Educación de Chile. (2009). *Currículum: Objetivos fundamentales y contenidos mínimos obligatorios de la educación básica y media*. Unidad de Currículum y Evaluación.
Ministerio de Educación de Chile. (2015). *Bases Curriculares 7° básico a 2° medio*. Unidad de Currículum y Evaluación.
Ministerio de Educación de Chile. (2016). *Programa de Estudio 2° medio, Matemática.* Unidad de Currículum y Evaluación.
Ministerio de Educación de Chile. (2018). *Bases Curriculares 1° a 6° básico*. Unidad de Currículo y Evaluación.
Ministerio de Educación de Chile. (2019a). *Bases Curriculares 3º y 4º medio.* Unidad de Currículum y Evaluación.
Ministerio de Educación de Chile. (2019b). *Programa de estudio matemática 3° medio para formación general*. Unidad de Currículum y Evaluación.

Ministerio de Educación de Chile. (2019c). *Programa de estudio matemática 4° medio para formación general.* Unidad de Currículum y Evaluación.
Ministerio de Educación de Chile. (2021a). *Programa de Estudio 3° medio, Matemática.* Unidad de Currículum y Evaluación.
Ministerio de Educación de Chile. (2021b). *Programa de Estudio 4° medio, Matemática.* Unidad de Currículum y Evaluación.
Ministerio de Educación de Chile. (2023). Actualización de la priorización curricular para la reactivación integral de aprendizajes: orientaciones generales. Unidad de Currículum y Evaluación.
Ministerio de Educación, Cultura y Deporte. (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato.* Boletín Oficial del Estado.
Morales, P. (2009). *Análisis de ítems en las pruebas objetivas.* Universidad Pontificia Comillas.
Morales, F., Vega, M., Aguilar, M. Gúmera, C. y Marchant, P. (2008). *Texto manual matemática 3º y 4º medio – proyecto explor@ndo*. Ediciones SM.
Muñoz, G., Gutiérrez, V. y Muñoz, S. (2013). T*exto matemática cuarto año medio.* Santillana.
Nasser, F. (2004). Structural Model of the Effects of Cognitive and Affective Factors on the Achievement of Arabic-Speaking Pre-service Teachers in Introductory Statistics. *Journal of Statistics Education, 12*(1), 1-19.
National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.
National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standars for Mathematics*. NGACBP and CCSSO.
Niess, M. (2005). Preparing teachers to teach science and mathematics with technology: Developing a technology pedagogical content knowledge. *Teaching and Teacher Education*, *21*(5), 509 523.
Noguera, M. (2017). *Evaluación de la idoneidad didáctica de una experiencia de enseñanza sobre distribuciones binomial y normal en 2º de bachillerato* [Tesis de Máster, Universidad de Granada].
Norambuena, P., Osorio, G., Romante, M. y Díaz, J. (2019). *Cuaderno de actividades de matemática 3° y 4° medio*. Ediciones SM.
Occeli, M. y Valeiras, N. (2013). Los libros de texto de ciencias como objeto de investigación: una revisión bibliográfica. *Enseñanza de las Ciencias*, *31*(2), 133- 152.
Ortiz, J. (2002). *La probabilidad en los libros de texto*. Grupo de Investigación en Educación Estadística de la Universidad de Granada.
Osorio, G., Norambuena, P., Romante, M., Gaete, D., Díaz, J., Celedón, J., Morales, K., Ortiz, N., Ramírez, P., Barrera, R. y Hurtado, Y. (2019). *Texto del estudiante de matemática 3° y 4° medio*. Ediciones SM.
Peñacoba, A., Muñoz, A.I. y Lafuente, M. B. (2022). Dimensiones latentes del proceso formativo del profesorado de Educación Secundaria: percepciones del alumnado. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 25*(3), 221-237.

Pérez, B. y Parraguez, M. (2013). Construcciones mentales de los conceptos aleatorio y determinista partir de la regresión lineal. En R. Flores (Ed.), *Acta Latinoamericana de Matemática Educativa, 26* (pp. 589–598). CLAME.

Persson, I., Kraus, K., Hansson, L. y Wallentin, F.Y. (2019). Confirming the structure of the survey of attitudes toward statistics(sats-36) by swedish students. *Statistics Education Research Journal, 18*(1), 83–93.

Pfannkuch, M. (2018). International handbook of research in statistics education. En D. Ben-Zvi, K. Makar y J. Garfield (Eds.), *Reimagining curriculum approaches* (pp. 387-413). Springer

Prado, M. y Gravoso, R. (2011). Improving high school students' statistical reasoning skills: a case of applying anchored instruction. *The Asia-Pacific Education Researcher, 20*(1), 61-72.

Ramírez, N. (28 de diciembre de 2018). *Siempre el 16% obtiene más de 600 puntos en la PSU: Demre explica cómo funciona la prueba. Emol*. https://www.emol.com/noticias/Nacional/2018/12/28/932495/Siempre-el-16-obtiene-mas-de-600-puntos-en-la-PSU-DEMRE-explica-como-funciona-la-prueba.html.

Rivas, H. (2014). *Idoneidad didáctica de procesos de formación estadística de profesores de educación primaria* [Tesis de doctorado no publicada]. Universidad de Granada.

Roa, R. (2000). *Razonamiento combinatorio en estudiantes con preparación matemática avanzada* [Tesis de Doctorado, Universidad de Granada].

Rodríguez-Santero, J. y Gil-Flores, J. (2019). Actitudes hacia la estadística en estudiantes de ciencias de la educación. Propiedades psicométricas de la versión española del Survey of Attitudes Toward Statistics (SATS-36). *Relieve*, *25*(1),1-17.

Rossman, A. (2008). *Reasoning about informal statistical inference: one statistician's view*. *Statistics Education Research Journal*, *7*(2), 5-19.

Ruiz, B. (2006). *Un acercamiento cognitivo y epistemológico a la didáctica del concepto de variable aleatoria* [Tesis de Máster, Instituto Politécnico Nacional].

Ruiz, B. (2013). *Análisis epistemológico de la variable aleatoria y comprensión de objetos matemáticos relacionados por estudiantes universitarios* [Tesis de Doctorado, Universidad de Granada].

Ruiz, K. (2019). *Comprensión del muestreo por alumnos chilenos de educación secundaria* [Tesis de Doctorado, Universidad de Granada].

Ruiz, B. y Albert, J. (2005). Didáctica de la Probabilidad y Estadística El Caso de la Variable Aleatoria. En J. Lezama, M. Sánchez y J. Molina (Eds.), *Acta Latinoamericana de Matemática Educativa, 18* (pp. 185- 191). CLAME.

Ruiz, B., Albert, J. y Batanero, C. (2006). An exploratory study of students' difficulties with random variable. En A. Rossman y B. Chance (Eds.), *7 International Conference on Teaching Statistics* (pp. 1-6). IASE.

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*(39), 431–449.

Salazar, R. (2014). *La variable aleatoria con probabilidad desde la perspectiva de la teoría APOE* [Tesis de Máster, Pontificia Universidad Católica de Valparaíso].

Salinas, J., Valdez, J. y Salinas-Hernández, U. (2018). Un acercamiento a la metodología lesson study para la enseñanza de la distribución normal. En L. Rodríguez-Muñiz, L. Muñiz-Rodríguez, A. Aguilar-González, P. Alonso, F. García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 525- 534). SEIEM.

Salinas-Herrera, J. y Salinas-Hernández, U. (2022). Teaching and Learning the Notion of Normal Distribution Using a Digital Resource. *Canadian Journal of Science, Mathematics and Technology Education, 22*(2), 1-15.

Sánchez, E. y Carrasco, G. (2018). El razonamiento probabilístico de estudiantes de bachillerato en actividades de distribución binomial. En L. Rodríguez-Múniz, L. Múñiz-Rodríguez, Á. Aguilar-González, P. Alonso, F. García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 535- 543). SEIEM.

Sánchez, E., Carrasco, G. y Herrera, M. (2018). Fundamental ideas in the probabilistic reasoning of high -school students in binomial distribution activities. En M. Sorto, A. White, y L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics* (pp.1-6). IASE.

Sánchez, E. y Landín, P. (2011). Fiabilidad de una jerarquía para evaluar el razonamiento probabilístico acerca de la distribución binomial. En M. Marín, G. Fernández, L. Blanco y M. Palarea (Eds.), *Investigación en Educación Matemática XV* (pp. 533- 541). SEIEM.

Sánchez, E. y Landín, P. (2014). Levels of probabilistic reasoning of high school students about binomial problems. En E. Chernoff y B. Sriraman (Eds.), *Probabilistic thinking presenting plural perspectives* (pp. 581-597). Springer.

Schulte, L. (2008). The development and validation of a teacher preparation program follow-up survey. *Journal of Statistics Education*, *16*(3), 1-10.

Setiawan, E. (2020). Introducing statistical inference to senior high school students: a textbook analysis. *Journal of Physics: Conference Series, 1663*, 1-9.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*(1), 1-36.

Sharma, S. (2006). High school students interpreting tables and graphs: implications for research. *International Journal of Science and Mathematics Education, 4*(2), 241-268.

Shin, B. M. (2012). A Study on a Didactic Transposition Method and Students' Understanding for the Normal Distribution. *Journal of Educational Research in Mathematics*, *22*(2), 117-136.

Soong, T. (2004). Fundamentals of probability and statistics for engineers. John Wiley and Sons.

Stahl, S. (2006). *The Evolution of the Normal Distribution. Mathematics Magazine*, *79*(2), 96-113.

Steinbring, H. (2005). *The construction of new mathematical knowledge in classroom interaction. An epistemological perspective*. Springer.

Stephenson, P., Richardson, M., Gabrosek, J. y Reischman. D. (2009). How LO can you GO? Using the Dice-Based Golf Game GOLO to Illustrate Inferences on Proportions and Discrete Probability Distributions. *Journal of Statistics Education, 17*(2), 1-37.

Stigler, J. y Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* The Free Press.

Tauber, L. (2001). *La construcción del significado de la distribución normal a partir de actividades de análisis de datos* [Tesis de Doctorado, Universidad de Sevilla].

Tauber, L., Sánchez, V. y Batanero, C. (2004). Diseño, implementación y análisis de una secuencia de enseñanza de la distribución normal en un curso universitario. *Revista EMA, 9*(2), 82-105.

Taufiq, I., Sulistyowati, F. y Usman, A. (2020). Binomial distribution at high school: an analysis based on learning trajectory. Journal of Physics: Conference Series, 1521(3), 1-7.

Toledo, Á., Montenegro, D. y Vicencio, I. (2019). Niveles de razonamiento frente a problemas binomiales. *Brazilian Journal of Development, 5*(6), 5399- 5410.

Torres, O. y Ojeda, A. (2018). Requisitos conceptuales de la función de densidad normal como modelo de la realidad. En D. García e I. Pérez (Eds.), *Acta Latinoamericana de Matemática Educativa, 31(2),* (pp. 1085-1093). CLAME.

Trouche, L., Gueudet, G. y Pepin, B. (2020). Documentational approach to didactics. En S. Lerman (Ed.), Encyclopedia of mathematics educations (pp. 237-247). Springer.

Tversky, A. y Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*(2), 105–110.

Tversky, A. y Kahneman, D. (1980). Causal schemas in judgments under uncertainty. En E. Fishbein (Ed.), *Progress in Social Psychology* (pp.49-72). Psychology Press.

Valdez, J. y Salinas, J. (2019). Análisis de las respuestas de estudiantes de bachillerato a problemas sobre la distribución normal. En J. M. Contreras, M. Gea, M. López-Martín y E. Molina-Portillo (Eds.*), Actas del Tercer Congreso Internacional Virtual de Educación Estadística* (pp. 1-10). Departamento de didáctica de la matemática de la Universidad de Granada.

Valverde, M. (2017). *Un estudio de la presentación de la distribución normal en los textos de bachillerato* [Tesis de Máster, Universidad de Granada].

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D. y Verschaffel, L. (2003). The illusion of linearity: Expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics, 53,* 113–138.

Vargas, V., Reyes, A. y Escalante, C. (2016). Ciclos de entendimiento de los conceptos de función y variación. *Educación matemática, 28*(2), 59-83.

Vásquez, C. y Alsina, A. (2015). Un modelo para el análisis de objetos matemáticos en libros de texto chilenos: situaciones problemáticas, lenguaje y conceptos sobre probabilidad. *Profesorado, revista de currículum y formación de profesorado, 19*(2), 441-462.

Vásquez, C. y Alsina. Á. (2017). Proposiciones, procedimientos y argumentos sobre probabilidad en libros de texto chilenos de educación primaria. *Profesorado, revista de currículum y formación de profesorado, 21*(1), 433-457.

Vilca, M. (2015). Tipificación de los errores que se presentan al identificar una variable aleatoria de distribución binomial en problemas contextualizados [Tesis de Máster, Pontificia Universidad Católica del Perú].

Williams, B., Onsman, A. y Brown, T. (2010). Exploratory factor analysis: a five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3), 1-13.

Wroughton, J. y Cole, T. (2013). Distinguishing Between Binomial, Hypergeometric and Negative Binomial Distributions. *Journal of Statistics Education, 21*(1), 1-16.

Yuan, Z. y Li, S. (2012). *Developing prospective mathematics teachers' technological pedagogical content knowledge (TPACK): a case of normal distribution* [Tesis de Doctorado, East China Normal University].

Zapico, M. (2006). Interrogantes acerca de análisis de contenido y del discurso en los textos escolares. En MINEDUC (Ed.), *Primer seminario internacional de textos escolares (pp.149-155)*. Ministerio de Educación de Chile.

Zieffler, A., Garfield, J., delMas, R. y Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 7*(2), 40-58.

Anexo 1: Significado institucional de variable aleatoria y distribuciones binomial y normal

1.1 Significado institucional de referencia de variable aleatoria y distribuciones binomial y normal según el currículo escolar chileno

Tabla A1

Significado institucional de referencia de variable aleatoria según el currículo escolar chileno

Situaciones-problemas			
S-P$_{1.1}$ Diferenciar entre variables aleatorias y variables con dependencia funcional			
S-P$_{1.2}$ Definir variables aleatorias finitas en experimentos aleatorios			
S-P$_{1.3}$ Identificar dominio de una variable aleatoria finita.			
S-P$_{1.4}$ Identificar recorrido o valores de una variable aleatoria finita			
S-P$_{2.1}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial			
S-P$_{2.2}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico			
S-P$_{2.3}$ Definir la función de probabilidad de una variable aleatoria utilizando la terminología $P(X=x)$			
S-P$_{2.4}$ Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta			
S-P$_{2.5}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta			
S-P$_{3.1}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua			
S-P$_{3.2}$ Calcular algunas probabilidades asociadas a una variable aleatoria continua.			
S-P$_{3.3}$ Determinar si la función dada es una función de probabilidad de una variable aleatoria continua			
S-P$_{4.1}$ Determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta			
S-P$_{4.2}$ Definir la función de distribución de una variable aleatoria discreta			
S-P$_{4.3}$ Representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta			
S-P$_{5.1}$ Calcular la media o valor esperado de una variable aleatoria discreta			
S-P$_{5.2}$ Interpretar la media o valor esperado de una variable aleatoria discreta			
S-P$_{5.3}$ Calcular la varianza de una variable aleatoria discreta			
S-P$_{5.4}$ Calcular la desviación estándar de una variable aleatoria discreta			
S-P$_{5.5}$ Interpretar la desviación estándar de una variable aleatoria discreta			
S-P$_{5.6}$ Identificar la media y desviación estándar de una variable aleatoria continua			
Objeto matemático	**10°**	**11°**	**12°**
Lenguaje			
L$_1$ Verbal	X	X	
L$_2$ Simbólico	X	X	
L$_3$ Gráfico	X	X	
L$_4$ Tabular	X	X	
Conceptos			
C$_1$ Variable aleatoria	X	X	
C$_2$ Función de probabilidad	X	X	
C$_3$ Función de densidad			X
C$_4$ Función de distribución	X		
C$_5$ Valores de posición central o de dispersión asociados a una variable aleatoria		X	
C$_{5.1}$ Media			
C$_{5.2}$ Varianza		X	
C$_{5.3}$ Desviación estándar		X	
Proposiciones			
PP$_1$ Caracterización de variable aleatoria y variables con dependencia funcional	X		
PP$_2$ Caracterización de variables aleatorias discreta y continua	X		
PP$_3$ Caracterización de la variable aleatoria mediante su función de probabilidad	X	X	
PP$_4$ Propiedades de la función de probabilidad	X	X	
PP$_6$ Propiedades de la función de densidad			X

Procedimientos			
P_1 Cálculo de la cardinalidad del dominio de una variable aleatoria discreta mediante combinatoria	X		
P_2 Partición disjunta del espacio muestral	X	X	
P_3 Determinación del dominio de una variable aleatoria discreta o probabilidad asociada utilizando diagrama de árbol	X		
P_4 Cálculo de probabilidades de una variable aleatoria discreta aplicando su función de probabilidad	X	X	
P_6 Cálculo de probabilidad de una variable aleatoria continua aplicando su función de densidad			X
P_7 Cálculo de valores de posición central y de dispersión asociados a una variable aleatoria discreta empleando sus expresiones algebraicas		X	
Argumentos			
A_1 Justificación por medio de ejemplo o contraejemplo	X	X	
A_2 Justificación a través de la representación gráfica	X		
$A_{3.1}$ Justificación mediante simulación con herramienta tecnológica	X		
$A_{3.2}$ Justificación mediante simulación con objetos manipulables	X		
A_5 Razonamiento algebraico	X	X	

Nota. Información extraída de Bizet et al. (2023b, 2023c).

Tabla A2

Significado institucional de referencia de distribuciones binomial y normal según el currículo escolar chileno

Situaciones-problemas		Situaciones-problemas	
S-$P_{6.1}$ Identificar situaciones que pueden modelizarse a través de una distribución binomial		S-$P_{7.1}$ Identificar situaciones que pueden modelizarse a través de una distribución normal	
		S-$P_{7.2}$ Determinar probabilidades asociadas a una distribución normal de forma manual	
S-$P_{6.2}$ Determinar los parámetros asociados a una distribución binomial		S-$P_{7.3}$ Describir la tendencia de los datos representados gráficamente empleando la aproximación de la binomial por la normal	
S-$P_{6.3}$ Calcular la media, varianza y desviación estándar de una distribución binomial		S-$P_{7.4}$ Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal	
S-$P_{6.4}$ Determinar probabilidades asociadas a una distribución binomial de forma manual		S-$P_{7.5}$ Determinar probabilidades asociadas a una distribución normal empleando una herramienta tecnológica	
S-$P_{6.5}$ Determinar probabilidades asociadas a una distribución binomial empleando una herramienta tecnológica		S-$P_{8.1}$ Calcular los parámetros asociados a una distribución normal	
		S-$P_{8.2}$ Calcular probabilidades en el contexto de aproximación de la distribución binomial a la normal	
Objeto matemático	**12°**	**Objeto matemático**	**12°**
Lenguaje		**Lenguaje**	
L_1 Verbal	X	L_1 Verbal	X
L_2 Simbólico	X	L_2 Simbólico	X
L_3 Gráfico	X	L_3 Gráfico	X
		L_4 Tabular	X
Conceptos		**Conceptos**	
C_1 Variable aleatoria	X	C_1 Variable aleatoria	X
C_6 Distribución binomial	X	C_7 Distribución normal	X
$C_{6.1}$ Función de probabilidad	X	$C_{7.1}$ Función de densidad	X
$C_{6.2}$ Función de distribución	X	$C_{7.2}$ Media	X
$C_{6.3}$ Media	X	$C_{7.3}$ Desviación estándar	X
		$C_{7.4}$ Moda	X

$C_{6.4}$ Varianza	X	$C_{7.5}$ Mediana	X
$C_{6.5}$ Desviación estándar	X	$C_{7.6}$ Función de densidad de la normal estándar	X
Proposiciones		Proposiciones	
PP_2 Caracterización de variables aleatorias discreta y continua	X	PP_2 Caracterización de variables aleatorias discreta y continua	X
PP_7 Caracterización de la distribución binomial	X	PP_8 Caracterización de la distribución normal	X
		PP_9 Propiedades de la función de densidad de la normal	X
		PP_{10} Propiedades para el cálculo de probabilidades con distribución normal estándar	X
		PP_{11} Propiedad de los intervalos centrales ($\pm 3\sigma$)	X
		PP_{12} Condición para aproximar una distribución binomial a una normal	X
Procedimientos		Procedimientos	
P_8 Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial	X	P_{11} Cálculo de probabilidad empleando la propiedad de los intervalos centrales ($\pm 3\sigma$)	X
P_9 Cálculo de probabilidades aplicando la función de distribución de una distribución binomial	X	P_{12} Tipificación	X
P_{10} Cálculo de media, varianza y desviación estándar de una binomial mediante sus expresiones algebraicas	X	P_{13} Cálculo de probabilidades utilizando la tabla de distribución normal estándar	X
Argumentos		Argumentos	
A_2 Justificación a través de la representación gráfica	X	A_2 Justificación a través de la representación gráfica	X
$A_{3.1}$ Justificación mediante simulación con herramienta tecnológica	X	$A_{3.1}$ Justificación mediante simulación con herramienta tecnológica	X
$A_{3.2}$ Justificación mediante simulación con objeto manipulable	X	A_4 Justificación mediante generalización	X
A_4 Justificación mediante generalización	X	A_5 Razonamiento algebraico	X
A_5 Razonamiento algebraico	X		

Nota. Información extraída de Bizet et al. (2023b, 2023c).

1.2 Significado institucional pretendido de variable aleatoria y distribuciones binomial y normal según los libros de texto analizados

Tabla A3

Configuración epistémica que define el significado institucional pretendido de variable aleatoria en los libros de texto analizados

Objeto matemático	10°	11°	12°
Situaciones-problemas			
S-$P_{1.1}$ Diferenciar entre variables aleatorias y variables con dependencia funcional	X		
S-$P_{1.2}$ Definir variables aleatorias finitas en experimentos aleatorios	X	X	X
S-$P_{1.3}$ Identificar dominio de una variable aleatoria finita.	X	X	X
S-$P_{1.4}$ Identificar recorrido o valores de una variable aleatoria finita	X	X	X
S-$P_{1.5}$ Diferenciar entre variables aleatorias discretas y continuas	X		X
S-$P_{2.1}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque frecuencial	X	X	
S-$P_{2.2}$ Determinar las probabilidades asociadas a los valores de una variable aleatoria discreta desde un enfoque clásico	X	X	X

S-P$_{2.3}$ Definir la función de probabilidad de una variable aleatoria utilizando la terminología P(X=x)	X	X	X
S-P$_{2.4}$ Representar en lenguaje tabular la función de probabilidad de una variable aleatoria discreta	X		
S-P$_{2.5}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria discreta	X	X	X
S-P$_{2.6}$ Calcular el valor de incógnitas para que la función propuesta satisfaga las condiciones de una función de probabilidad	X	X	
S-P$_{3.1}$ Representar en lenguaje gráfico la función de probabilidad de una variable aleatoria continua			X
S-P$_{3.2}$ Calcular algunas probabilidades asociadas a una variable aleatoria continua			X
S-P$_{3.3}$ Determinar si la función dada es una función de probabilidad de una variable aleatoria continua			X
S-P$_{3.4}$ Calcular el valor de una incógnita para que la función dada sea de probabilidad			X
S-P$_{4.1}$ Determinar la probabilidad acumulada de algunos valores de una variable aleatoria discreta	X	X	
S-P$_{4.2}$ Definir la función de distribución de una variable aleatoria discreta	X	X	
S-P$_{4.3}$ Representar en lenguaje gráfico la función de distribución de una variable aleatoria discreta	X	X	
S-P$_{5.1}$ Calcular la media o valor esperado de una variable aleatoria discreta		X	X
S-P$_{5.2}$ Interpretar la media o valor esperado de una variable aleatoria discreta		X	X
S-P$_{5.3}$ Calcular la varianza de una variable aleatoria discreta		X	X
S-P$_{5.4}$ Calcular la desviación estándar de una variable aleatoria discreta		X	X
S-P$_{5.5}$ Interpretar la desviación estándar de una variable aleatoria discreta		X	X
S-P$_{5.6}$ Identificar la media y desviación estándar de una variable aleatoria continua			X
Lenguaje			
L$_1$ Verbal	X	X	X
L$_2$ Simbólico	X	X	X
L$_3$ Gráfico	X	X	X
L$_4$ Tabular	X	X	X
Conceptos			
C$_1$ Variable aleatoria	X	X	X
C$_2$ Función de probabilidad	X	X	
C$_3$ Función de densidad			X
C$_4$ Función de distribución	X	X	
C$_5$ Valores de posición central o de dispersión asociados a una variable aleatoria			
C$_{5.1}$ Media		X	X
C$_{5.2}$ Varianza		X	X
C$_{5.3}$ Desviación estándar		X	X
Proposiciones			
PP$_1$ Caracterización de variable aleatoria y variables con dependencia funcional	X		
PP$_2$ Caracterización de variables aleatorias discreta y continua	X		X
PP$_3$ Caracterización de la variable aleatoria mediante su función de probabilidad	X	X	X
PP$_4$ Propiedades de la función de probabilidad	X	X	
PP$_5$ Caracterización de la variable aleatoria mediante su función de distribución	X	X	
PP$_6$ Propiedades de la función de densidad			X
Procedimientos			
P$_1$ Cálculo de la cardinalidad del dominio de una variable aleatoria discreta mediante combinatoria	X		
P$_2$ Partición disjunta del espacio muestral	X	X	
P$_3$ Determinación del dominio de una variable aleatoria discreta o probabilidad asociada a sus valores utilizando diagrama de árbol	X		
P$_4$ Cálculo de probabilidades de una variable aleatoria discreta aplicando su función de probabilidad	X	X	X
P$_5$ Cálculo de probabilidad de una variable aleatoria discreta aplicando su función de distribución	X	X	

P_6 Cálculo de probabilidad de una variable aleatoria continua aplicando su función de densidad			X
P_7 Cálculo de valores de posición central y de dispersión asociados a una variable aleatoria discreta empleando sus expresiones algebraicas		X	X
Argumentos			
A_1 Justificación por medio de ejemplo o contraejemplo	X	X	X
A_2 Justificación a través de la representación gráfica	X	X	X
$A_{3.1}$ Justificación mediante simulación con herramienta tecnológica	X	X	
A_5 Razonamiento algebraico	X	X	X
A_6 Razonamiento verbal-deductivo	X	X	
A_7 Justificación mediante análisis-síntesis	X	X	X

Nota. Información extraída de Bizet et al. (2023a, 2023b).

Tabla A4

Configuración epistémica que define el significado institucional pretendido de distribuciones binomial y normal en los libros de texto analizados

Objeto matemático			Objeto matemático	
Situaciones-problemas	11°	12°	Situaciones-problemas	12°
S-$P_{6.1}$ Identificar situaciones que pueden modelizarse a través de una distribución binomial	X	X	S-$P_{7.1}$ Identificar situaciones que pueden modelizarse a través de una distribución normal	X
			S-$P_{7.2}$ Determinar probabilidades asociadas a una distribución normal de forma manual	X
S-$P_{6.2}$ Determinar los parámetros asociados a una distribución binomial	X	X	S-$P_{7.3}$ Describir la tendencia de los datos representados gráficamente empleando la distribución normal o la aproximación de la binomial por la normal	X
			S-$P_{7.4}$ Evaluar la veracidad de afirmaciones utilizando el concepto de distribución normal	X
S-$P_{6.3}$ Calcular la media y varianza de una distribución binomial	X	X	S-$P_{7.6}$ Calcular la probabilidad en una distribución normal estándar	X
			S-$P_{7.7}$ Representar en lenguaje gráfico una distribución normal	X
			S-$P_{7.8}$ Calcular los valores correspondientes a una probabilidad dada	X
S-$P_{6.4}$ Determinar probabilidades asociadas a una distribución binomial de forma manual	X	X	S-$P_{8.1}$ Calcular los parámetros asociados a una distribución normal	X
			S-$P_{8.2}$ Calcular probabilidades en el contexto de aproximación de la distribución binomial a la normal	X
Lenguaje			Lenguaje	
L_1 Verbal	X	X	L_1 Verbal	X
L_2 Simbólico	X	X	L_2 Simbólico	X
L_3 Gráfico		X	L_3 Gráfico	X
			L_4 Tabular	X
Conceptos			Conceptos	
C_1 Variable aleatoria	X	X	C_1 Variable aleatoria	X
C_6 Distribución binomial	X	X	C_7 Distribución normal	X
$C_{6.1}$ Función de probabilidad	X	X	$C_{7.1}$ Función de densidad	X
$C_{6.2}$ Función de distribución	X		$C_{7.2}$ Media	X
$C_{6.3}$ Media		X	$C_{7.3}$ Desviación estándar	X
$C_{6.4}$ Varianza		X	$C_{7.6}$ Función de densidad de la normal estándar	X
Proposiciones			Proposiciones	
PP_2 Caracterización de variables aleatorias discreta y continua	X	X	PP_2 Caracterización de variables aleatorias discreta y continua	X

			PP_6 Propiedades de la función de densidad	X
			PP_8 Caracterización de la distribución normal	X
PP_7 Caracterización de la distribución binomial	X	X	PP_9 Propiedades de la función de densidad de la normal	X
			PP_{10} Propiedades para el cálculo de probabilidades con distribución normal estándar	X
			PP_{11} Propiedad de los intervalos centrales $(\pm 3\sigma)$	X
			PP_{12} Condición para aproximar una distribución binomial a una normal	X
Procedimientos			Procedimientos	
P_8 Cálculo de probabilidades aplicando la función de probabilidad de una distribución binomial	X	X	P_{11} Cálculo de probabilidad empleando la propiedad de los intervalos centrales $(\pm 3\sigma)$	X
			P_{12} Tipificación	X
P_9 Cálculo de probabilidades aplicando la función de distribución de una distribución binomial	X	X	P_{13} Cálculo de probabilidades utilizando la tabla de distribución normal estándar	X
P_{10} Cálculo de media y varianza de una binomial mediante sus expresiones algebraicas		X	P_{14} Corrección por continuidad	X
Argumentos			Argumentos	
A_1 Justificación por medio de ejemplo o contraejemplo	X	X	A_1 Justificación por medio de ejemplo o contraejemplo	X
A_2 Justificación a través de la representación gráfica		X	A_2 Justificación a través de la representación gráfica	X
			$A_{3.1}$ Justificación mediante simulación con herramienta tecnológica	X
A_5 Razonamiento algebraico	X	X	A_4 Justificación mediante generalización	X
			A_5 Razonamiento algebraico	X

Nota. Información extraída de Bizet et al. (2023a, 2023b).

Anexo 2: Conclusions

2.1 Introduction

This chapter presents the final conclusions of this doctoral thesis, organized in three sections. In the first one, a discussion on the achievement of the research objectives and hypotheses set out in Chapter 1 is presented. In section 6.3, a synthesis of the main contributions of the thesis to Statistics Education is shown. In the last section, the limitations of this research are analyzed and some lines for future work are presented.

2.2 Findings on the Research Objectives and Hypotheses

2.2.1 Conclusions on the first Aim and Hypothesis

(O1): *Conduct a literature review on the teaching and learning of random variables and binomial and normal distributions.*

To achieve this goal, a study presented in Chapter 2 has been carried out. In Study 1(Bizet et al., 2022), the leading research reported in the literature in the field of mathematics education between 1990 and 2020, on the teaching and learning of the random variable and its probability distribution, both at school and university level, was analyzed. The results showed that fewer studies have been found that deal with its epistemological and didactic aspects (teaching) compared to those that deal with its cognitive elements (learning).

Concerning the epistemological field, research has superficially identified elements of the random variable (problems, procedures, different representations, etc.) necessary for its teaching at school (Ortiz, 2002) and university (Ruiz, 2013). For the binomial model, no studies have been found that provide guidelines on the main elements to be considered in its teaching (although in the background complementary to Study 1, i.e., works published after 2020, one was found on its origin and historical development). Contrary to what happened with the normal, the results obtained are more encouraging, given that the fundamental elements to be included in its teaching process were found in detail, both in higher education (Tauber, 2001) and compulsory school education (Valverde, 2017).

Several papers on the didactic aspects of university and school education were found. These focus on proposals for teaching the random variable and the regular model at the university level and the binomial at the school level. At the same time, there are fewer instructional proposals in the school context on the random and normal variables and at the university level on the binomial.

Within the studies developed from a cognitive perspective, we found about the same amount of research that addresses the understanding of the random variable and its probability distribution in school and university education. However, gaps in the literature would be essential to investigate, for example, i) the level of understanding of the random variable and probability models at the end of school education or ii) the level of understanding of the binomial distribution in higher education. On the other hand, some difficulties have been identified that students at school and/or university level manifest when working with the random variable, the binomial distribution and/or the normal distribution. However, studies have yet to be found to interpret those difficulties or errors around the topics of interest regarding semiotic conflicts.

In this way, the results evidenced in Study 1 confirm the hypothesis that the analysis of previous research on the teaching and learning of random variables and/or binomial and normal distributions shows that they have been subjects of limited investigation. Some research was recognized about the analysis of research on the teaching of random variables and probability models at school. However, the research did not develop an in-depth investigation of the elements (problem situations, propositions, procedures, languages, etc.) involved in their tasks, nor did they analyze the problem fields that can be identified. About the investigation of research on the learning of subjects of interest at the school level, no studies have been found that analyze the understanding of random variables and probability models at the end of school. No studies were found to interpret difficulties or errors regarding semiotic conflicts.

2.2.2 Conclusions on the second Aim and Hypotheses

(O2): *Analyze the treatment of random variables and binomial and normal distributions in Chilean curricular regulations and textbooks.*

This objective was achieved by developing two studies (Study 2 and Study 3) presented in Chapter 3. In the first one (Bizet et al., 2023a), we analyzed the problem situations on random variables and binomial and normal distributions in a sample of five Chilean textbooks organized in two series, where each one covered grade 10 (15 to 16 years old), 11 (16 to 17 years old) and 12 (17 to 18 years old). Initially, based on the Guide of Problems-Situations on Random Variables and their applications in Probability Distributions according to the Chilean School Curriculum (GPS-RVPD: Bizet et al., 2023c), the activities present in the books surveyed were classified, recognizing the type(s) of situation-problems involved in each one. Subsequently, the problem situations were contrasted with the topics proposed in the Chilean school curriculum and textbooks.

The results of Study 2 have shown that, with regard to the diversity of problem-situations on the random variable, 24 types were observed in the books analyzed, organized in five problem fields: P-F_1 Identify the random variable as a function present in the probabilistic context; P-F_2 Recognize the probability distribution of a discrete random variable as a tool to see random variation; P-F_3 Recognize the probability distribution of a continuous random variable as a tool to see random variation; P-F_4 Establish the distribution function of a discrete random variable; and P-F_5 Use some values of central position or dispersion linked to the random variable.

Likewise, in the school textbooks surveyed, many problem situations on applying the random variable in probability distributions were evidenced. Particularly for the binomial model, four types were observed, structured in one problem field, P-F_6 The binomial distribution as a probabilistic model describing a real-life phenomenon. For the normal model 9 types were observed, organized in two problem fields: P-F_7 The normal distribution as a probabilistic model describing a real-life phenomenon; and P-F_8 Approximating distributions of discrete random variables with many values.

In this way, it was observed that there are situations promoted in the textbooks analyzed that were not suggested in the Chilean curriculum, such as differentiating between discrete and continuous random variables, calculating the value of unknowns so that the proposed function is a probability function (for both types of variables indicated), calculating probabilities in a standard normal distribution, representing the normal distribution in a graph, and calculating the values corresponding to a given probability in the context of the normal distribution. It was also evident that situations-problems identified in the Chilean curriculum were not addressed in school textbooks or had minimal presence, such as calculating probabilities associated with binomial and normal distributions using a technological tool and differentiating between random variables and variables with functional dependence, respectively. Therefore, the results demonstrate a need for more alignment between the MINEDUC (2015; 2019a) proposal and the textbooks analysed.

Study 3 (Bizet et al., 2023b) investigated the language, concepts, propositions, procedures and arguments related to the random variable and the binomial and normal distributions in the sample of five Chilean textbooks previously used (Study 2). The mathematical objects of interest were first recognized in the Chilean school curriculum, then identified in the sample of books, and finally, results were contrasted. Regarding the language related to the random variable, it was found that there is congruence between the curricular guidelines and the two series of books since four types of language have been proposed in each document. A scenario similar to the one described above was found for the normal distribution. In the context of the binomial, disharmony was observed between the documents surveyed since the curriculum suggested three types of language, as did one of the book series and the other series exclude graphic language.

Regarding the concepts for the random variable, coherence between the curricular guidelines and the two series of books was evident, as seven concepts were presented in each document. In probability models, discrepancies were observed between the documents analyzed. The curriculum promoted six concepts for binomial distribution, unlike the two-book series that omitted some. One omitted its distribution function and standard deviation, and the other omitted its mean, variance, and standard deviation. Typically, the curriculum suggested seven concepts. At the same time, the two series omitted some, one of them omitting the standard normal density function, mean, variance, mode and median, and the other omitting the last two.

Some discrepancies were identified among the documents surveyed on the propositions. Specifically on the random variable, the curriculum suggested five propositions, excluding its characterization using the distribution function and through the density function, while one series of books proposed seven and the other six, the latter omitting the characterization of a random variable and variables with functional dependence. Furthermore, for the normal, the curriculum guidelines proposed six propositions, discarding the properties of the density function, unlike each series, which suggested seven. However, in the context of the binomial, there was harmony between the documents analyzed since the curriculum suggested two propositions equal to those proposed in each series.

There was also disharmony in the procedures among the documents surveyed. The curriculum for the random variable promoted six procedures, ignoring calculating probabilities with its distribution function. At the same time, one set of books proposed seven and the other six, excluding determining probabilities or their dominance using tree diagrams. The curriculum guidelines presented three procedures for the binomial, as did one set of books. However, the other set ruled out calculating its mean, variance and standard deviation using algebraic expressions. For the normal domain, the curriculum suggested three procedures except for the continuity correction, while each set of books suggested four.

In addition, inconsistencies were found in the arguments between the documents analyzed. In the random variable, the curriculum guidelines promoted five types, suppressing verbal-deductive reasoning and justification through analysis-synthesis, unlike the two series of books: one promoted five, discarding simulations using technological tools and manipulable objects, and the other promoted six, omitting the last named. The curriculum presented five arguments for the binomial, excluding validation by example or counterexample. One set of books included three, excluding argumentation through generalization and simulations using a technological tool and manipulable object, and the other set presented only two, omitting those three and justification using graphs.

In the context of the normal, the curricular guidelines promoted four types of arguments, omitting justification by example or counterexample. At the same time, one series of books suggested three, discarding validation using simulation with a technological tool and generalization, and the other series contemplated four, suppressing justification using graphic representation. In this way, it was found that there are primary mathematical objects linked to the random variable and binomial and normal distributions that have been identified in the Chilean curriculum but were excluded from the textbooks, and vice versa, demonstrating a lack of coherence between the guidelines proposed by MINEDUC (2015; 2019a) and that resource.

Furthermore, the analyses developed in Study 2 and Study 3, presented in Chapter 3, have made it possible to establish i) the institutional reference meaning of random variable and binomial and normal distributions according to the Chilean school curriculum and ii) the intended institutional meaning of random variable and binomial and normal distributions according to the textbooks analysed.

Finally, the results of both Study 2 and Study 3 confirm the hypothesis that in the treatment of random variable and binomial and normal distributions, there are differences between that suggested by the Chilean curricular regulations and that proposed by the Chilean textbooks. On the one hand, discrepancies were observed in the problem situations (activities, tasks, exercises and problems) on the topics of interest suggested in the Chilean curricular guidelines and those promoted in the Chilean textbooks analysed. On the other hand, discrepancies were found between the documents analysed about i) the propositions, procedures and arguments related to the random variable; ii) the language, concepts, procedures and arguments related to the binomial distribution; and iii) the concepts, propositions, procedures and arguments related to the normal distribution.

2.2.3 Conclusions on the third Objective and Hypothesis

(O3): *To construct and validate an instrument aimed at assessing the understanding of random variables and binomial and normal distributions in the Chilean school context.*

This objective was achieved through two studies (Study 4 and Study 5) in Chapter 4. The first study dealt with the construction of the instrument and its content validity, for which a rigorous methodological process involving several phases was followed. In contrast, the second study presented the construct validity and reliability of the instrument.

In Study 4 (Bizet et al., 2023c), the Guide of Problem Situations on Random Variables and their Applications in Probability Distributions according to the Chilean School Curriculum (GPS-RVPD) was created, and the construction process was organized into two sub-stages. The first sub-stage dealt with identifying problem fields on the topics of interest based on an epistemological analysis. Specifically, based on the analysis of research on epistemological aspects of the random variable, binomial distribution, and normal distribution, set out in Chapter 2, and through a cyclical and inductive process, eight problem fields were established (P-F_1 to P-F_8) linked to the topics in question.

The second sub-stage for constructing the GPS-RVPD addressed recognising problem situations on the topics of interest from an analysis of curriculum guidelines. Specifically, based on a content analysis of the Chilean school curriculum, the American curriculum guidelines and the well-known GAISE project, current regulations (units of analysis) on the teaching-learning of the random variable and probability distributions were identified. The units of analysis were then classified into the eight problem domains (categories of analysis) previously defined. Finally, those units of analysis have been compared and reduced, and from these 33 problem situations on a random variable, binomial distribution and normal distribution are inferred, which together form the GPS-RVPD.

In the second phase of Study 4, an initial set of items was chosen from the literature to compose the assessment instrument. Specifically, for each problem situation of the GPS-RVPD, three items were proposed, which were selected from previous research, tasks proposed in previously analysed textbooks (Chapter 3) or elaborated by the authors. It should be noted that two problem situations (P-$S_{6.5}$ determine probabilities associated with a binomial distribution using a technological tool and P-$S_{7.5}$ determine probabilities associated with a normal distribution using a technological tool) were excluded at this stage because the instrument was intended to be applied to individual students and either face-to-face or online.

In the third phase, expert judgment chooses the final set of items that make up the instrument through content validity. Six PhDs in mathematics education with research experience in statistics didactics participated and independently evaluated three items for each task. The initial version of the instrument was obtained by selecting the item with the highest mean and lowest standard deviation in each problem situation. In addition, the Content Validity Coefficient (Hernández-Nieto, 2002) was calculated for each item in order to analyze the degree to which it was adjusted to a problem situation, and the instrument was found to have a good validity and concordance coefficient (0,87).

In Study 5 (Bizet et al., under evaluation), the initial version of the instrument (32 items) was administered to a pilot sample of 80 Chilean university students who had yet to take a course on statistics and probability at university. The instrument's construct validity was then tested by means of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), and its reliability was tested using Cronbach's alpha coefficient.

The results of the EFA analysis showed that the empirical structure of the questionnaire is composed of six factors that explain 58% of the total variance, each one related to a problem domain (consisting of three to four items) on the topics in question:(i) the first factor could be associated with P-F_4 Establish the distribution function of a discrete random variable; (ii) the second factor to P-F_2 Recognise the probability distribution of a discrete random variable as a tool that allows one to see random variation; (iii) the third factor to P-F_6 The binomial distribution as a probabilistic model describing a real-life phenomenon; (iv) the fourth factor to P-F_5 Use some values of central position or dispersion linked to the random variable; (v) the fifth factor to P-F_7 The normal distribution as a probabilistic model describing a real-life phenomenon; and (vi) the sixth factor possible to link to P-F_1 Identify the random variable as a function present in the probabilistic context.

Thus, in the empirical structure of the instrument, two problem fields were discarded: i) P-F_3 linked to the probability function of a continuous random variable. However, part of its content was included in P-F_1, using $I_{3.1}$ representing P-$S_{3.1}$ on graphing that function; and ii) P-F_8 related to the approximation of distributions of discrete random variables, although the approximation of the binomial by the normal was dealt with in P-F_7, using $I_{7.3}$ representative of P-$S_{7.3}$ on describing the trend of data represented graphically using that approximation.

The CFA results showed that most of the factor-item coefficients have adequate values. Regarding the factor-factor coefficients, it was found that there is an average relationship between the answers to P-F_2 linked to the probability function of a discrete random variable, with the answers to P-F_1 involving the random variable as a function, and with the answers to P-F_4 about the distribution function of a discrete random variable. In addition, there is a medium relationship between the problem domains on probabilistic models, i.e. between P-F_6 binomial distribution and P-F_7 normal distribution. While there is a low relationship between the answers: i) to P-F_1 linked to the random variable as a function and P-F_5 related to some values of central position or dispersion associated with the random variable; and ii) to P-F_2 on the probability function of a discrete random variable and P-F_6 related to the binomial distribution.

Also, in the CFA, the empirical structure of the instrument consisting of six factors and 19 items was found to have acceptable goodness-of-fit measures. On the other hand, Cronbach's alpha of the instrument was calculated, and the value obtained showed that the questionnaire is reliable and that there is consistency between the items and construct.

Study 3 and Study 4 were complemented by the calculation of the difficulty and discrimination index for each of the 19 items that make up the final version of the instrument. The results showed that most items had adequate discrimination; three were moderately complex, and 16 were difficult. Therefore, the results obtained in Chapter 4 guarantee that the final version of the questionnaire (19 items) is valid (it has content and construct validity) and reliable to investigate the understanding of random variables and binomial and normal distributions in Chilean school graduates. Furthermore, the designed questionnaire characterises both the assessed institutional meaning of random variables and binomial and normal distributions.

Thus, the results obtained in Chapter 4 confirm the hypothesis that the instrument designed to assess the understanding of random variables and binomial and normal distributions has good psychometric characteristics. In this sense, the adequate internal consistency of the questionnaire and its six-factor factor structure are highlighted.

2.2.4 Conclusions on the fourth Objective and Hypothesis

(O4): *To assess the understanding of random variables and binomial and normal distributions in Chilean school leavers.*

This objective was achieved through Study 6 (see Chapter 5), which assessed the understanding of random variables and probability models in 101 Chilean school leavers using a previously validated questionnaire (Chapter 4) composed of 19 items. The overall results of this study showed that the total score obtained by each participant varied between 0 and 20 points (out of a maximum score of 38 points), and the average score achieved was 4 points.

Regarding the understanding of random variables, the results showed that the participants solved most situations-problems assessed in the questionnaire, but in a low percentage. The tasks with the highest proportion of partial or correct answers were: i) P-$S_{1.1}$ differentiate between random variables and variables with functional dependence (43%) and P-$S_{3.1}$ represent in graphical language the density function of a continuous random variable (17,82%), belonging to P-F_1; and ii) P-$S_{2.2}$ determine the probabilities associated with the values of a discrete random variable from the classical probability approach (17,82%), which is part of P-F_2.

The tasks with the lowest percentage of partially correct or correct answers were: i) P-$S_{4.3}$ represent in graphical language the distribution function of a discrete random variable (1%), belonging to P-F_4; and ii) P-$S_{5.4}$ identify the mean and standard deviation of a continuous random variable (1%) and P-$S_{5.5}$ interpret the standard deviation of a discrete random variable (0%), belonging to P-F_5.

In addition, in the tasks on assessing the understanding of random variables, specific semiotic cognitive conflicts were observed, mainly in the partial and erroneous answers, such as: confusing the path with the domain of a random variable, assuming equiprobability of compound random events, confusing the probability function with the density function (or vice versa), confusing the mean with the mode of a discrete random variable, arguing an answer based on personal beliefs about the context of the task, among others.

In short, the personal meaning of the random variable achieved by the participants lacks particular primary mathematical objects that differentiate it from the institutional meaning assessed. For example, none of the young people managed to answer the problem of interpreting the standard deviation of a discrete random variable. Even more limited students in the sample have managed to use the different mathematical objects involved in solving the tasks assessed. Consequently, Chilean school leavers need a higher understanding of random variables.

Regarding the understanding of probability models, the results showed that the participants solved all the situations assessed in the questionnaire, although in a low proportion. The tasks with the highest percentage of partial or correct answers were: i) P-$S_{6.4}$ determine probabilities associated with a binomial distribution manually (15%), which is part of P-F_6; and ii) P-$S_{7.4}$ evaluate the veracity of statements using the concept of the binomial distribution (43,56%), which belongs to P-F_7. The tasks with the lowest percentage of partially correct or correct answers were: i) P-$S_{6.1}$ identify situations that can be modelled using a binomial distribution (1%) and P-$S_{6.3}$ calculate the mean of a binomial distribution (1%) belonging to P-F_6; and ii) P-$S_{7.2}$ determine probabilities associated with a normal distribution manually (1%) belonging to P-F_7.

Likewise, in the tasks related to assessing the understanding of probability models, some semiotic cognitive conflicts were observed, mainly in the partial and erroneous answers, for example: assuming equiprobability of compound random events, confusing the mean with the median of a discrete random variable, and translating from verbal to symbolic language incorrectly in the calculation of probabilities, among others.

In summary, the personal meaning of binomial distribution achieved by the students in the sample is consistent with the institutional meaning assessed in the questionnaire. However, few participants used the diversity of primary mathematical objects associated with the resolution of the assessed problem-situations. Therefore, the participants need a better understanding of binomial distribution.

Finally, the personal meaning of normal distribution achieved by the young participants lacks primary mathematical objects that differentiate it from the assessed institutional meaning, such as: tabular language; the concepts of mean, standard deviation, symmetry and standard normal distribution; the property of central intervals, those for calculating probabilities with standard normal distribution and the conditions for approximating a binomial by a normal; and the process of calculating probabilities, typing and reading the standard normal table. Also, a few students in the sample applied some primary objects associated with the solution of the assessed tasks. Therefore, the participants need a better understanding of the normal distribution.

Thus, the results obtained in Chapter 5 confirm the hypothesis that the participants need an understanding of random variables and binomial and normal distributions. In this sense, the results evidenced in each problem area stand out. In P-F_1, identifying the random variable as a function present in the probabilistic context, one participant achieved the maximum score (6 points), although the average score achieved was 1,2 points. For P-F_2, recognizing the probability distribution of a discrete random variable as a tool that allows us to see random variation, no student reached the maximum score (8 points), and the average score obtained was 1,2 points. Also, in P-F_4, to establish the distribution function of a discrete random variable, and in P-F_5, to use some values of central position or dispersion linked to the random variable, none of the participants achieved the maximum score (6 points), and the average score achieved was 0,1 and 0,2 points respectively.

Furthermore, for P-F_6, the binomial distribution as a probabilistic model describing a real-life phenomenon, one participant achieved the maximum score (6 points), but the mean score was 0,7. Finally, for P-F_7, the normal distribution as a probabilistic model describing a real-life phenomenon, none of the students achieved the maximum score (6 points), and the mean score obtained was 0,5.

2.3 Main contributions

The analysis of curricular guidelines, mainly from Chile and the United States, has allowed us to identify the main tasks on random variables and binomial and normal distributions suggested at the school level. Thus, a Guide of Situations-Problems on Random Variables and their Applications in Probability Distributions according to the School Curriculum (Bizet et al., 2023c) has been proposed, which has valuable information that can be useful for mathematics teachers working in schools to guide their work and design teaching proposals on the topics in question.

The studies of Chilean textbooks have made it possible to expand the types of activities on random variables and normal model present in previous research around this didactic resource (Ortiz, 2002; Valverde, 2017).In addition, they made it possible to complement the concepts, languages, procedures, propositions and arguments related to the random variable and normal distribution recognized in isolation in previous studies of school texts (Doukhan and Guedet, 2019; Ortiz, 2002; Valverde, 2017). Another novel contribution was the identification of tasks on the binomial model and various primary mathematical objects involved in its solution, promoted by school texts since no studies have been found that address the treatment of the binomial in this didactic resource, differentiating the various mathematical objects involved.

Thus, studies on textbooks (Bizet et al., 2023a; 2023b) can be valuable input for teachers who must teach the topics of random variables and/or binomial and normal distributions in schools. Also, those studies can have valuable information for textbook authors since they provide criteria to improve the elaboration of textbooks, which continue to be the most used didactic resource to teach mathematics in school education (Vásquez and Alsina, 2015).

On the other hand, we have contributed with a validated questionnaire to evaluate the understanding of random variables and binomial and normal distributions in school graduates (Bizet et al., under evaluation) based on curricular guidelines and the Ontosemiotic Approach to Mathematical Knowledge and Instruction. In addition, this instrument could have several uses: i) to evaluate the understanding of school students after a teaching-learning process of random variables and/or probability models; ii) to diagnose the understanding of students who have just entered university on these topics; or iii) to assess the mathematical knowledge of random variables and binomial and normal models in future mathematics teachers, since these are topics that they will have to teach at school.

As for the comprehension evaluation study, it has contributed to establishing the level of comprehension of the random variable and binomial and normal distributions of Chilean school graduates based on the EOS. This study has also contributed to establishing the personal meaning of the topics of interest achieved by the participants and the semiotic conflicts that manifest when working with the random variable and probability models. This information is original, mainly so it could contribute to developing statistics education.

2.4 Limitations and future lines of research

Concerning the limitations of this research, some have been identified in its different phases. In the studies on the analysis of Chilean textbooks, one restriction has been the number of series used, so if those series were extended, a better characterization of the treatment of random variables and binomial and normal distributions in this critical resource could be obtained. Another limitation of the work was the omission of the investigation of the activity booklet resource that each textbook has since not all of them were available.

In the study on the instrument's construct validity, one of the limitations was the size of the sample; due to the health emergency caused by the COVID-19 pandemic, it took much work to gain access to Chilean first-year university students. Another of the drawbacks of this study was how to implement the instrument, since as a result of Covid-19, the questionnaire had to be applied online during a 90-minute session. In this sense, a complementary study using a more prominent and representative sample and implementing the questionnaire in person would be necessary to confirm or refute the results obtained in Study 5.

In the study on comprehension assessment, one of the restrictions was the type of sample (non-probabilistic). This has been chosen intentionally according to the availability of Chilean universities that participated in the study, where, due to the Covid-19 pandemic, it was difficult for professors to give up an online session of their subject, and students agreed to answer an online questionnaire of open questions and send photographs of their work. Another area for improvement was how the instrument was implemented. As the questionnaire was applied online, most participants who responded to the tasks omitted their justification, although it was requested.

Some future lines of research are projected below. Specifically in textbooks, it would be desirable to identify potential semiotic conflicts around treating the random variable and binomial and normal distributions.

On the other hand, the evaluation study results pose the challenge of conducting research that addresses the design and implementation of a didactic unit aimed at school students from 16 to 18 years of age that gradually introduces the problem fields of random variables and probability models. Subsequently, to analyze if the implementation of the designed questionnaire generates an improvement in the understanding of the topics.

Finally, it would be interesting to investigate the understanding of random variables and binomial and normal distributions in future mathematics teachers, including mathematical and didactic knowledge. Because these are proposed in curricular guidelines, they should be taught in schools soon.