



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO
DEPARTAMENTO DE
MATEMÁTICA EDUCATIVA

**“Validación de un marco de desarrollo cognitivo de las concepciones de
nociones básicas de muestreo estadístico de estudiantes universitarios y
profesores de bachillerato”**

T E S I S

Que presenta:

FELIPE DE JESÚS JIMÉNEZ RODRÍGUEZ

Para obtener el grado de

MAESTRO EN CIENCIAS

EN LA ESPECIALIDAD EN
MATEMÁTICA EDUCATIVA

Director de tesis:

DR. ERNESTO ALONSO SÁNCHEZ SÁNCHEZ

Ciudad de México

Agosto, 2024

Agradezco al CONAHCYT, por el apoyo económico brindado para realizar mis estudios de Maestría.

Becario 1080804

Agradecimientos

Quiero expresar mi más profundo agradecimiento a mi director de tesis, el Dr. Ernesto Sánchez Sánchez, por su invaluable guía y apoyo durante todo este proceso. Sin su experiencia y dedicación, este trabajo no habría sido posible.

A los miembros del jurado de tesis, Dra. Guadalupe Carrasco Licea y Dra. Ana María Ojeda Salazar, por su apoyo y disposición que han tenido hacia mi trabajo.

A Francisco Sepúlveda por sus comentarios, sugerencias y el apoyo con la exploración con los profesores. También a José Luis López, por su asistencia técnica y apoyo constante.

Agradezco a todo el personal del Departamento de Matemática Educativa, en particular a Adriana Parra por su apoyo permanente y eficaz.

A mi mamá, Norma, por darme todo su amor. A mis hermanos, Ana, María José, Bruno y Martha; por su cariño. A mi padre, Felipe, por su apoyo constante. A mi abuelo, por ser ejemplo de fortaleza. A mis suegros, Minerva y Roberto, por su cariño y apoyo.

Finalmente, agradezco a Pris, mi compañera, por todos los años de aprendizaje y crecimiento. Agradezco por tu paciencia, especialmente durante mis momentos de estrés y duda. Gracias a ti he aprendido a ver el mundo con otros ojos y a desafiar mis límites, sobre todo gracias por ser ejemplo de constancia y valentía para lograr lo que uno se propone. También, a Milos, el gato, por sacarme tantas sonrisas con su cálida compañía.

Tabla de contenidos

Agradecimientos	3
Tabla de contenidos	4
Resumen	6
Capítulo I. Introducción y planteamiento del problema.....	8
Estructura del documento	11
Capítulo II. Antecedentes.....	13
Jerarquía o progresión de aprendizaje y marco de desarrollo cognitivo.....	13
Marcos de desarrollo en la Educación Estadística	15
Capítulo III. Marco teórico.....	25
Marco de desarrollo cognitivo sobre concepciones	25
Marco de desarrollo cognitivo para algunos conceptos de muestreo	28
Sobre la validación de un marco de desarrollo	32
Dimensiones de validación	32
Capítulo IV. Método	37
Entrevista a una profesora especializada en estadística	38
Primera exploración con el instrumento de toma de datos	40
Valoración de expertos en educación estadística	42
Aplicación del cuestionario a dos Modelos de Lenguaje con Inteligencia Artificial	44
Segunda exploración con profesores de CCH	46
Capítulo V. Presentación de resultados.....	49
Entrevista a una Especialista en Estadística	49
Primera exploración con el instrumento	52
Evaluación del Cuestionario por Expertos.....	57
Aplicación del Cuestionario a dos Modelos de Lenguaje de Inteligencia Artificial.....	62
Segunda exploración con el instrumento con profesores de CCH.....	65
Capítulo VI. Conclusiones	69
La exploración con profesores	69
El uso de GPT-3.5 y GPT-4.....	70
La consulta con expertos.....	70
La exploración con estudiantes.....	70
La entrevista con la maestra.....	71
Acciones para investigaciones futuras:.....	71

Referencias	72
Apéndice A: Instrumento presentado en la entrevista con la profesora experta en estadística.	77
Apéndice B: cuestionario usado en la exploración con alumnos.....	80
Apéndice C: cuestionario sobre concepciones acerca de conceptos de muestreo	84
Apéndice D: cuestionario aplicado a los modelos de inteligencia artificial.....	92
Apéndice E: cuestionario aplicado a profesores de CCH	98
Apéndice F: modificaciones realizadas en cada fase de implementación.....	104

Resumen

Este trabajo tiene como objetivo principal validar un marco de desarrollo cognitivo relacionado con las concepciones de población, muestra, muestreo repetido y distribución muestral en estudiantes universitarios y profesores de bachillerato. Se busca identificar qué recursos o instrumentos son efectivos para ofrecer evidencia que contribuya a la validez de este marco de desarrollo. Se implementaron cinco etapas de recolección de datos: (1) entrevistas con una académica universitaria experta en la materia, (2) aplicación de un instrumento diseñado a cuatro estudiantes universitarios, (3) evaluación y perspectiva de tres expertos en investigación de educación estadística, (4) experimentación con modelos de lenguaje GPT-3.5 y GPT-4 utilizando el instrumento, y (5) aplicación de un cuestionario a once docentes de bachillerato. Los resultados mostraron la utilidad del instrumento para discernir distintos niveles de abstracción en las concepciones de los docentes y la tendencia de los estudiantes a responder en un nivel práctico aun después de haber llevado un curso de estadística. Los hallazgos sugieren que el marco de desarrollo cognitivo propuesto es efectivo para reflejar los procesos de conceptualización de los estudiantes y docentes en los temas de población, muestra, muestreo repetido y distribución muestral.

Palabras clave: Desarrollo cognitivo, concepciones, muestreo, distribución muestral, educación estadística, validación de instrumentos.

Abstract

The main objective of this work is to validate a cognitive development framework related to the conceptions of population, sample, repeated sampling, and sampling distribution in university students and high school teachers. The aim is to identify which resources or instruments are effective in providing evidence that contributes to the validity of this development framework. Five stages of data collection were implemented: (1) interviews with a university academic expert in the field, (2) application of a designed instrument to four university students, (3) evaluation and perspective of three experts in statistical education research, (4) experimentation with GPT-3.5 and GPT-4 language models using the instrument, and (5) application of a questionnaire to eleven high school teachers. The results showed the usefulness of the instrument in discerning different levels of abstraction in teachers' conceptions and the tendency of students to respond at a practical level even after taking a statistics course. The findings suggest that the proposed cognitive development framework is effective in reflecting the conceptualization processes of students and teachers in the areas of population, sample, repeated sampling, and sampling distribution.

Capítulo I. Introducción y planteamiento del problema

El propósito principal de este estudio es la presentación y validación de un marco de desarrollo cognitivo de las concepciones de población, muestra, muestreo repetido y distribución muestral que desarrollan los estudiantes universitarios y profesores de bachillerato durante sus estudios de la estadística.

Un marco de desarrollo es una descripción o modelo de cómo progresa y crece la comprensión o el aprendizaje de los estudiantes en un tema determinado, tal descripción se realiza a través de alguna manifestación de la comprensión o el aprendizaje como: desempeño, razonamiento, pensamiento o concepciones, mismos que se observan durante la realización de una tarea apropiada. Un marco de desarrollo cognitivo es una herramienta para diversos fines educativos, por ejemplo, puede ser útil en la estructuración del currículo sobre el tema, para hacer un diagnóstico o evaluación de las habilidades o conocimientos de los estudiantes y como auxiliar del profesor para el diseño de sus planes de clase. Además, un marco de desarrollo cognitivo puede ser la base para posteriores investigaciones, ya sea para refinar y profundizar el conocimiento didáctico sobre el tema, ya para coadyuvar en la exploración de nuevas estrategias de enseñanza.

Por otro lado, el corazón de la estadística es la inferencia y, por tanto, es una parte sustancial de los cursos de estadística. La inferencia permite hacer afirmaciones probables sobre una población con base en la información que proporciona una muestra. Para hacer inferencias se requiere de un conjunto de conceptos y técnicas que se agrupan bajo el término general de muestreo. Estrictamente, el muestreo es el proceso para obtener muestras de una población, pero para entender, describir y justificar tal proceso se definen y utilizan muchos términos, de los cuales hemos escogido para el marco de desarrollo los de población, muestra, muestreo repetido y distribución muestral.

También se ha elegido poner atención en las *concepciones* que los estudiantes se forman acerca de tales términos porque es con base en ellas que se pueden explicar sus respuestas a las tareas y los razonamientos que las sustentan. De acuerdo con Confrey (1990) una concepción es una creencia, teoría, significado o explicación que los estudiantes desarrollan en relación con un término matemático o científico. Una concepción puede ser

una forma embrionaria, ingenua, parcial o inmadura de un concepto, pero conteniendo aspectos esenciales para una futura comprensión más madura. Por otro lado, una concepción también puede ser diferente o equivocada respecto al concepto que designa el término, en este caso se llama mala concepción o concepción errónea. Nuestro marco trata de captar la evolución de las concepciones de las estudiantes alineadas parcialmente con el significado normativo del concepto.

La elaboración de marcos de desarrollo para el muestreo en diferentes niveles escolares ha sido de interés para los investigadores en educación matemática. Por ejemplo, Watson y Moritz (2000b) propusieron un marco de desarrollo de las ideas de estudiantes de grados 3, 6 y 9 (edades entre 9 y 15 años) sobre el término “muestra”, de cómo lo aplican en situaciones en contexto y de su actitud crítica hacia expresiones que lo incluyen. Aunque Saldanha y Thompson (2002) no presentan formalmente un marco de desarrollo, sí que lo prefiguran. En efecto, informan que los estudiantes de bachillerato tienen dos concepciones de muestra distintas y progresivas: aditiva, que considera una muestra simplemente como un subconjunto de la población, y multiplicativa, que percibe la muestra como una versión proporcional a pequeña escala de la población. Estas concepciones constituyen el comienzo de un posible marco sobre el concepto de muestra para estudiantes de bachillerato. Finalmente, Case y Jacobbe (2018) sugieren un marco de dos ejes para caracterizar las dificultades de los estudiantes: Un eje se refiere a la dificultad de distinguir y comprender la distribución de la población, la de una muestra y la distribución muestral; el otro se refiere a la dificultad de los estudiantes para distinguir una perspectiva real de una perspectiva hipotética.

Langrall et al. (2017) sugieren que “se necesitan estudios que contrasten y validen estos marcos conceptuales o de desarrollo...” refiriéndose a los marcos sobre los conceptos de variación y distribución que se han propuesto recientemente por diferentes investigadores, parece razonable trasladar esta recomendación para los marcos que se refieren al muestreo, pero agregando la sugerencia de que se amplíen o refinen. El problema de cómo validar los marcos de desarrollo de conceptos estadísticos no se ha elaborado suficientemente en la literatura de educación estadística.

Jones et al. (1997) está entre los pocos estudios que presentan y se proponen validar un marco sobre el pensamiento probabilístico de niños de primaria. El proceso de validación que llevaron a cabo consistió en traducir las descripciones del marco en tareas e incluirlas en un protocolo de entrevista para recoger datos de los pensamientos de los estudiantes y con estos: 1) Refinar los descriptores del marco, 2) Verificar la consistencia a través de los constructos y 3) Iluminar (sic) las características de cada nivel dentro del marco. Los autores aclaran que utilizan métodos cualitativos para analizar datos y apuntalar la validación. Más recientemente Lobato y Walters (2017) señalan que el método más común para validar una jerarquía, que es similar a un marco de desarrollo, es el uso de la teoría de respuesta al ítem (IRT) basada en modelos de Rasch; en educación estadística lo utilizaron Watson et al. (2003) para validar una jerarquía sobre el concepto de variación. Este método es de tipo cuantitativo y su eficacia depende de la calidad del instrumento de medición y de la obtención de un gran número de datos.

En el presente estudio seguimos una línea más cercana al análisis cualitativo, afín a la propuesta de Jones et al. (1997), sin renunciar a considerar aspectos cuantitativos básicos, sin embargo, no tuvimos la oportunidad de contar con una población de estudiantes amplia de modo que fue imposible utilizar la teoría de respuesta al ítem ni la técnica Rasch.

Dada la importancia de elaborar y validar marcos conceptuales sobre el muestreo en educación estadística y teniendo en cuenta la dificultad de implementar procesos para validarlos, en este trabajo nos proponemos presentar un marco para describir y modelar las concepciones sobre los constructos de población, muestra, muestreo repetido y distribución muestral para estudiantes universitarios y profesores de bachillerato. Para esto formulamos las siguientes preguntas de investigación:

- ¿Qué recursos o instrumentos, bajo las restricciones del presente estudio, son útiles para ofrecer evidencia que contribuya a la validez del marco de desarrollo sobre conceptos de muestreo?
- ¿Qué nos dicen los datos que resultan de la implementación de tales recursos e instrumentos acerca de la validez o no del marco?

En los siguientes capítulos aclaremos cuáles son los instrumentos utilizados también ampliaremos y discutiremos el concepto de validación del marco. Además, presentaremos

los datos que surgieron a partir de los instrumentos creados y discutiremos los resultados del análisis de los datos.

Estructura del documento

El documento se encuentra organizado en seis capítulos. En el primer capítulo, Introducción y planteamiento del problema, se establece el propósito del estudio, que es validar un marco de desarrollo cognitivo sobre las concepciones de población, muestra, muestreo repetido y distribución muestral en estudiantes universitarios y profesores de bachillerato. Se introduce el concepto de marco de desarrollo, explicando su importancia para la estructura curricular, evaluación, y se subraya la relevancia de la inferencia en la estadística.

En el segundo capítulo, Antecedentes, se presentan las diferentes acepciones de las jerarquías y progresiones de aprendizaje en educación matemática, con un enfoque en probabilidad y estadística. Se exploran estudios sobre el razonamiento estadístico y los conceptos de población, muestra, muestreo repetido y distribución muestral, mencionando tres tipos de jerarquías: desarrollo, secuencia didáctica y contenido.

En el tercer capítulo, Marco teórico, se detalla el aparato conceptual que sustenta el trabajo, incluyendo los componentes del contenido estadístico y los mecanismos cognitivos involucrados en el tratamiento de la distribución muestral, así como el proceso de validación del marco y el enfoque utilizado.

En el cuarto capítulo, Método, se describe el proceso metodológico del estudio, las características de los participantes, los instrumentos y datos recolectados, y las condiciones de implementación de las trayectorias de aprendizaje. Se especifican cinco fases de recolección de datos, incluyendo entrevistas con expertos y experimentación con modelos de lenguaje de inteligencia artificial.

En el quinto capítulo, Presentación de resultados, se presentan los hallazgos de las cinco etapas de recolección de datos, destacando la utilidad del instrumento para discernir niveles de abstracción en las concepciones de los docentes y sugiriendo la efectividad del marco de desarrollo cognitivo propuesto.

En el sexto capítulo, Conclusiones, se resumen las observaciones y hallazgos principales, discutiendo las implicaciones de los resultados y proponiendo mejoras para futuros estudios, concluyendo que el marco de desarrollo refleja efectivamente los procesos de conceptualización en los temas de población, muestra, muestreo repetido y distribución muestral, y sugiriendo acciones para investigaciones futuras.

Capítulo II. Antecedentes

Nuestra investigación se centra en la validación de un marco de desarrollo, este capítulo presenta brevemente las diferentes acepciones de las jerarquías y progresiones de aprendizaje en educación matemática, con un enfoque particular en los campos de probabilidad y estadística. Se examinan estudios que exploran el razonamiento estadístico y que se relacionan con los conceptos de población, muestra, muestreo repetido, distribución muestral, los cuales son analizados en esta investigación.

Jerarquía o progresión de aprendizaje y marco de desarrollo cognitivo

En sus primeras manifestaciones una jerarquía en Educación Matemática se refería a cómo y en qué orden los niños aprenden matemáticas, pero se entendía de varias maneras. Siguiendo a Hart (1981) se podían distinguir tres tipos:

- (i) *Jerarquía de desarrollo*: Descripción de formas sucesivas en que los individuos conciben, piensan, comprenden o razonan un tema matemático.
- (ii) *Secuencia didáctica*. Presentaciones, actividades y/o problemas ordenados para implementar sucesivamente en el aula y propiciar el aprendizaje del tema
- (iii) *Contenido*. Una exposición lógica de conceptos, procedimientos y problemas que desarrollan el tema matemático.

Las *jerarquías de desarrollo* pretendían identificar rasgos de la conducta o de la mente de los individuos respecto a un tema matemático en estrecha relación con los niveles de desarrollo cognitivo del aprendiz, típicamente basados en las etapas de Piaget, aunque algunas otras estaban basadas en modelos alternativos (Carpenter, 1979).

En los últimos 50 años ha habido una abundante investigación sobre trayectorias o progresiones de aprendizaje en educación matemática. Lobato y Walters (2017) hicieron una reseña/estudio de los enfoques de las trayectorias y/o progresiones de aprendizaje publicadas en educación matemática o educación en ciencias. Parten de la definición general de una progresión de aprendizaje del National Research Council (2007):

Las progresiones de aprendizaje son descripciones de formas sucesivamente más sofisticadas de pensar sobre un tema que pueden sucederse a medida que

los niños aprenden e investigan un tema durante un amplio período de tiempo (por ejemplo, de 6 a 8 años). (p. 219)

No obstante, Lobato y Walters (2017) encuentran una variedad amplia de enfoques. En general, las jerarquías/progresiones de aprendizaje se consideran como herramientas potenciales para alinear el currículo, los estándares, la evaluación y el desarrollo profesional, reflejando un interés creciente en cómo se puede guiar mejor el aprendizaje estudiantil a través de descripciones detalladas de formas cada vez más sofisticadas de pensar en un tema a lo largo del tiempo.

Estos autores descubren en la literatura siete enfoques diferentes adoptados por investigadores que identifican su trabajo como una trayectoria o progresión de aprendizaje; ellos los nombran enfoques de:

1. Niveles cognitivos
2. Niveles de discurso
3. Esquemas y operaciones
4. Trayectorias hipotéticas de aprendizaje
5. Prácticas matemáticas colectivas
6. Lógica disciplinar y coherencia curricular
7. Estrategias observables y rendimientos de aprendizaje.

El presente trabajo se ubica dentro de los estudios que proponen trayectorias/progresión del enfoque de *Nivel cognitivo* y adoptamos la expresión *Marco de desarrollo cognitivo* para referirnos a este tipo de jerarquía o progresión; esta decisión nos acerca un poco a la expresión de *Marco conceptual y de desarrollo* que propone Langrall et al. (2007) para referirse a las trayectorias propuestas en el campo de educación estadística.

En un *Marco de desarrollo cognitivo* se describen estados sucesivos de algún tipo de cognición (v. gr. comprensión, concepciones, formas de razonamiento o de pensamiento) acerca de un contenido matemático o estadístico específico.

Modelo SOLO

Expondremos muy brevemente el modelo SOLO porque ha influido en la construcción de la mayoría de las jerarquías o marcos que se han construido en investigaciones de diferentes contenidos estadísticos. En efecto, en la última década del siglo pasado comenzó a influir en el campo de la educación estadística y probabilística la propuesta teórica neo-piagetiana conocida como SOLO (Structure of the Observed Learning Outcomes) (Biggs y Collis, 1982, 1991). De acuerdo con Biggs y Collis (1991) en el desarrollo cognitivo humano se pueden identificar cinco *Modos de Funcionamiento*, similares a las etapas de Piaget: 1) *Sensoriomotor*, 2) *Icónico*, 3) *Simbólico-concreto*, 4) *Formal* y 5) *Post-formal*. Además, dentro de cada modo de funcionamiento se producen ciclos de aprendizaje que tienen cinco niveles jerárquicos: 1) *Prestructural*, 2) *Uniestructural*, 3) *Multiestructural*, 4) *Relacional* y 5) *Abstracto extendido*. Cada uno de estos niveles sirve para caracterizar la calidad de las respuestas que los sujetos desarrollan frente a una tarea compleja cuya solución implica la articulación de dos o más componentes. Es muy importante aclarar que estos niveles no son para clasificar a los sujetos, sino a sus respuestas a la tarea. Una respuesta se clasifica en el nivel Prestructural cuando es irrelevante con relación a la solución normativa de la tarea. Es Uniestructural cuando refiere a un solo componente de la solución. Es Multiestructural cuando se consideran dos o más componentes de la solución, pero se utilizan de manera independiente una de la otra sin relacionarlas entre sí. Es relacional cuando se consideran las componentes y sus relaciones de manera que la solución integra todas las componentes. Finalmente, el nivel Abstracto extendido se produce cuando el estudiante generaliza la estructura revelada por la tarea a otras tareas diferentes con una estructura de solución similar. Este nivel no se puede reconocer en la respuesta a la tarea dada, sino que requiere observaciones adicionales.

Marcos de desarrollo en la Educación Estadística

Entre los primeros trabajos que propusieron marcos de desarrollo en la investigación de la educación estadística se encuentra un modelo simple propuesto por Shaughnessy (1992) para el concepto de probabilidad. Su cobertura va desde una total carencia de comprensión de los eventos aleatorios hasta una habilidad para comparar y contrastar varias versiones

matemáticas de ellos; sus niveles son: 1) No estadístico, 2) Ingenuo, 3) Emergente, y 4) Pragmático.

Jones et al. (1997) desarrollaron “Un marco para evaluar y nutrir el pensamiento probabilístico de niños pequeños [Tercer grado de primaria]” teniendo en cuenta la jerarquía SOLO. El contenido del pensamiento probabilístico lo caracterizan con cuatro constructos: *Espacio muestral*, *Probabilidad de un evento*, *Comparación de probabilidades* y *Probabilidad condicional*. Los niveles que proponen son: 1) Subjetivo, 2) Transicional, 3) Cuantitativo informal y 4) Numérico. Por lo tanto, organizan los descriptores en una tabla/matriz constructo vs nivel. Tarr y Jones (1997) proponen un marco para evaluar el pensamiento de los estudiantes de nivel secundaria (middle school) sobre los conceptos de probabilidad condicional e independencia estocástica. Los autores proponen cuatro niveles para estos conceptos con los mismos nombres de Jones et al. (1997). Jones et al. (2000) diseñaron un marco de desarrollo para caracterizar el pensamiento estadístico de los niños (del 1° al 5° grado), también teniendo como base la jerarquía SOLO. Utilizan cuatro constructos característicos del pensamiento estadístico: 1. *Descripción de datos*. 2. *Organización y reducción de datos*. 3. *Representación de datos*. 4. *Análisis e interpretación de datos*. Los niveles que propusieron son: 1. *Idiosincrático*, 2. *Transicional*, 3. *Cuantitativo*, y 4 *Analítico*.

Con la introducción del análisis exploratorio de datos, los avances tecnológicos y búsqueda de cambiar las prácticas educativas existentes, la investigación en educación estadística ha experimentado un notable desarrollo, orientándose hacia una perspectiva más integral del pensamiento estadístico y probabilístico (Langrall et al., 2017). Este enfoque se centra especialmente en la inferencia informal y en la comprensión profunda de la variabilidad, que a su vez deriva en el análisis de conceptos como muestra, muestreo repetido, distribución, distribución muestral, etc. Los hallazgos de las investigaciones orientadas con este enfoque sugieren que los estudiantes muestran distintos niveles de razonamiento en estos temas, y a su vez, ha impulsado el desarrollo de marcos diseñados para facilitar la observación y el análisis de la evolución del pensamiento estadístico en los estudiantes. En tabla 2.1 presentamos un breve panorama de algunos marcos que se describen las concepciones de estudiantes y que sirven de antecedentes para este trabajo.

Tabla 2.1*Clasificación de las investigaciones de acuerdo a los conceptos en que hacen énfasis*

Población	Muestra	Muestreo Repetido	Distribución Muestral
	Watson y Moritz (2000b)		
	Saldanha y Thompson (2002)		
			Chance et al. (2004)
		Saldanha y Thompson (2007)	
		Noll & Shaugnessy (2012)	
		van Dijke-Droogers et al. (2019)	
	Silvestre et al. (2022)		

A continuación, detallaremos aspectos de investigaciones que describen concepciones estadísticas de estudiantes.

Concepto de muestra como punto de partida del razonamiento inferencial

En su estudio Watson y Moritz (2000b), analizan cómo estudiantes de distintos niveles educativos, tercer, sexto, noveno y undécimo grado; comprenden y construyen el concepto de muestra en estadística a través de la aplicación de un cuestionario y una entrevista a los participantes. La investigación examina tres aspectos del razonamiento sobre muestreo:

- Método de recolección: el método empleado para seleccionar muestras.
- Tamaño de la muestra: el tamaño adecuado de estas
- Sesgo en el muestreo: la identificación y manejo de los sesgos potenciales.

Los hallazgos permitieron a los autores elaborar un marco descriptivo que busca categorizar el desarrollo conceptual del muestreo en seis niveles distintos (Tabla 2.1). Estos niveles sugieren una evolución en la comprensión y habilidades de los estudiantes: desde un enfoque en tamaños de muestra pequeños, típicamente inferiores a 15, hasta el reconocimiento de la necesidad de muestras más grandes en niveles avanzados; de la adopción de métodos de

selección basados en criterios personales o poco sistemáticos hacia la implementación de métodos jerarquizados o aleatorios más rigurosos; finalmente, de una inicial falta de atención o comprensión de los sesgos a un aumento progresivo en la conciencia de su relevancia y de las estrategias para mitigarlos.

Tabla 2.2

Marco de desarrollo del concepto de muestra de Watson y Moritz (2000a)

Pequeñas muestras* sin selección.

- puede proporcionar ejemplos de muestras, como productos alimenticios
- puede describir una muestra como una pequeña parte o, más raramente, como un intento o prueba
- aceptar un tamaño de muestra de menos de 15
- no sugerir ningún método de selección o un método idiosincrásico

Pequeñas muestras con selección aleatoria primitiva.

- proporcionar ejemplos de muestras, como productos alimenticios
- describir una muestra como una pequeña parte o un intento o prueba
- aceptar un tamaño de muestra de menos de 15
- sugerir selección aleatoria sin descripción o con una simple instrucción elegir "cualquiera", tal vez de diferentes escuelas

Pequeñas muestras con selección de Resultados

- proporcionar ejemplos de muestras, como productos alimenticios
- describir una muestra como un pequeño fragmento y un intento o prueba
- aceptar un tamaño de muestra de menos de 15
- sugerir la selección de personas por peso, ya sea personas delgadas o de peso normal

Muestras grandes con selección aleatoria o distribuida

- proporcionar ejemplos de muestras, como productos alimenticios
- describir una muestra como una pequeña parte y como un intento o prueba puede referirse al término promedio
- sugerir un tamaño de muestra de al menos 20 o un porcentaje de la población
- sugerir una selección basada en un proceso aleatorio o distribución por geografía

Muestras grandes sensibles al sesgo

- proporcionar ejemplos de muestras, que a veces implican encuestas
- describir una muestra como un pequeño fragmento y un intento o prueba
- puede referirse a términos promedio o representativos
- sugerir un tamaño de muestra de al menos 20 o un porcentaje de la población
- sugerir una selección basada en un proceso aleatorio o distribución por geografía
- expresar preocupación por la selección de muestras para evitar sesgos
- identificar muestras sesgadas en artículos periodísticos que informan sobre los resultados de las encuestas

Muestras equívocas

- proporcionar ejemplos y descripciones de muestras
- puede indicar diferencias sobre el tamaño de la muestra, a veces basándose en aspectos irrelevantes
- pueden basar sus decisiones en un tamaño pequeño con métodos de selección apropiados o con sensibilidad parcial al sesgo, o basar sus decisiones en un tamaño de muestra grande con métodos de selección inapropiados

Nota. *En la versión original se utiliza el término “sampler” (muestreador), pero en español no se usa éste.

Por su parte, Saldanha y Thompson (2002) en un experimento de enseñanza identificaron dos maneras distintas en que en estudiantes 11° y 12° concebían las *muestras*.

Al pedirles a los estudiantes que usaran un software para simular muestreos repetidos y visualizar la distribución muestral, los investigadores clasificaron en dos categorías la forma en que los participantes del estudio concebían las muestras. La primera, que denominaron concepción aditiva, se refiere a la idea de una muestra simplemente como una colección de individuos seleccionados de una población más grande, sin una relación proporcional necesaria con la población total. La segunda, conocida como concepción multiplicativa, sugiere una comprensión más compleja, en la que los estudiantes ven la muestra no sólo como un subconjunto sino también como una representación a pequeña escala de la población, reflejando las proporciones y características de la población de manera casi proporcional.

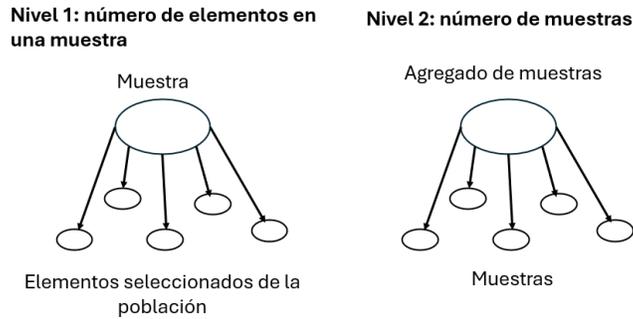
De las dos investigaciones anteriores podemos ver que Watson y Moritz (2000) enfocan su análisis en los aspectos normativos del muestreo, examinando cómo los estudiantes evolucionan en sus estrategias de muestreo a medida que adquieren conocimiento. En contraste, Saldanha y Thompson (2002) se centran en las diferencias entre dos visiones distintas del muestreo: la aditiva y la multiplicativa. Estas concepciones influyen en la forma en que los estudiantes racionalizan las muestras, cuestión que veremos más adelante, influye en la forma que conceptualizan la distribución muestral.

El muestreo repetido rumbo la construcción de la distribución muestral

Saldanha y Thompson (2007) exploraron cómo los estudiantes de bachillerato realizan inferencias estadísticas informales basadas en la comprensión de las distribuciones muestrales, tanto empíricas —formadas por objetos en urnas— como simuladas —generadas mediante software. En su experimento de enseñanza, analizan cómo los estudiantes realizan inferencias e interpretan frecuencias de un atributo obtenidos de muestras. En su análisis rastrearon el cambio conceptual desde una visión singular, donde las muestras se consideran de manera aislada, hacia una visión distribucional, donde los estudiantes operan con colecciones de muestras observando la variabilidad de éstas.

Figura 2.1

Niveles de razonamiento en torno a las muestras y muestreo repetido (Saldanha y Thompson (2007)).



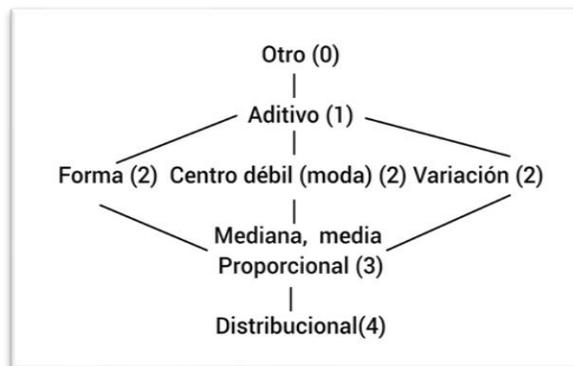
Por su parte, Noll y Shaughnessy (2012) reportaron cómo estudiantes de sexto a duodécimo grado conceptualizan y anticipan los resultados de la distribución muestral empírica, específicamente en el contexto de muestreo repetido. Para ello, realizaron una actividad práctica donde los participantes estimaban la proporción de elementos de una urna al realizar muestras de ésta.

A partir de los resultados, propusieron un marco teórico de cuatro niveles jerárquicos de comprensión de la distribución muestral:

1. Aditivo: Los estudiantes simplemente suman los resultados sin considerar las propiedades de la distribución.
2. Forma, Centro y Variación: Reconocen características como la media o el rango de los datos.
3. Proporcional: Consideran las proporciones de los diferentes colores en su muestreo.
4. Distribucional: Comprensión madura que incluye el reconocimiento de la distribución de los datos como un todo y la variabilidad asociada con el muestreo.

Figura 2.2

Rejilla de evolución del concepto distribución muestral (Noll & Shaughnessy, 2012).



La distribución muestral como concepto complejo

Chance et al. (2004) examinaron las dificultades de los estudiantes universitarios para comprender las distribuciones muestrales. Afirman que aprender sobre distribuciones muestrales es desafiante para los estudiantes, ya que requiere integrar múltiples conceptos estadísticos y razonar sobre muestras hipotéticas. Los autores proponen que la instrucción debe abordar conceptos previos como *distribución, muestra, población, variabilidad y muestreo*. Además, sugieren métodos como la exploración práctica del muestreo, el uso de simulaciones, y la discusión en clase de las observaciones hechas por los estudiantes. Los autores destacan que, aunque el uso de software de simulación es una herramienta útil, por sí solo no es suficiente para superar los conceptos erróneos de los estudiantes sobre las distribuciones de muestreo.

Tabla 2.3

Conocimientos previos necesarios para aprender sobre distribuciones muestrales, (Chance et al., 2004)

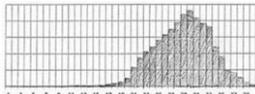
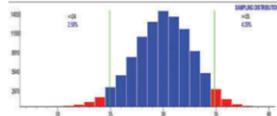
-
- *La idea de variabilidad.* ¿Qué es una variable? ¿Qué significa decir que las observaciones varían? Los estudiantes necesitan una comprensión de la extensión de una distribución en contraste con los conceptos erróneos comunes de uniformidad o variedad.
 - *La idea de una distribución.* Los estudiantes deben poder leer e interpretar presentaciones gráficas de datos cuantitativos y describir el patrón general de variación. Esto incluye ser capaz de describir distribuciones de datos; caracterizando su forma, centro y extensión; y poder comparar diferentes distribuciones sobre estas características. Los estudiantes deben poder ver entre los datos y describir la forma general de la distribución, y estar familiarizados con las formas comunes de las distribuciones, como normal, sesgada, uniforme y bimodal.
 - *La distribución normal.* Esto incluye propiedades de la distribución normal y cómo una distribución normal puede verse diferente debido a cambios en la variabilidad y el centro. Los estudiantes también deben estar familiarizados con la idea del área bajo una curva de densidad y cómo el área representa la probabilidad de los resultados.
 - *La idea de muestreo.* Esto incluye muestras aleatorias y cómo son representativas de la población. Los estudiantes deben sentirse cómodos distinguiendo entre una muestra estadística y un parámetro de población. Los estudiantes deberían haber comenzado a considerar o ser capaces de considerar cómo las estadísticas de muestra varían de una muestra a otra, pero siguen un patrón predecible.
-

Por su parte, van Dijke-Droogers et al. (2019) realizaron una trayectoria de aprendizaje estructurada en tres etapas para enseñar a estudiantes de noveno grado a realizar inferencias estadísticas. La primera etapa introdujo a los estudiantes a los conceptos de

muestra y muestreo a través de tareas que implicaban estimar la cantidad de bolas amarillas en una caja opaca. En la segunda etapa, los estudiantes aprendieron sobre la variación muestral y el tamaño de la muestra mediante simulaciones de muestreo. Finalmente, en la tercera etapa, profundizaron en el muestreo repetido y utilizaron distribuciones muestrales simuladas para entender cómo la variabilidad de los datos informa sobre inferencias estadísticas precisas de una población desconocida.

Tabla 2.4

Actividades en cada paso en la trayectoria de aprendizaje (van Dijke-Droogers et al. (2019))

	Paso 1	Paso 2	Paso 3
Actividad de enseñanza	Realizar un experimento con la caja negra física (con ventana pequeña y grande)	Imaginar la distribución de frecuencia de más de 100,000 repeticiones del experimento de caja negra física	Simulación de la distribución de muestral de una caja negra física con TIC (Tecnologías de Información y Comunicación)
Concepto	Muestra Variación de muestras Muestreo repetido Tamaño de la muestra	Distribución de frecuencia de los datos de muestreo repetido	Distribución de muestral simulada del muestreo repetido
Razonamiento Inferencial	Argumentación sobre la variación de muestreo y el tamaño de la muestra	Argumentación sobre la distribución de frecuencia basada en resultados de muestras (im)probables	Razonamiento inferencial con la distribución de muestreo simulada sobre variación e incertidumbre
Actividad del estudiante	Estimar el tamaño y el contenido de la caja negra 	Esbozar la distribución de frecuencia esperada a partir de datos de más de 100,000 repeticiones, determinando resultados de muestras (im)probables 	Razonamiento inferencial con la distribución de muestreo simulada 

Los autores mencionan que la secuenciación de las etapas de la trayectoria de aprendizaje fue significativa, ya que facilitó a los estudiantes sin experiencia previa en muestreo a razonar con datos de muestra en poco tiempo, incluyendo el manejo de variación e incertidumbre. Añaden que, los estudiantes pudieron realizar inferencias más precisas sobre la población mediante el enfoque de muestreo repetido para minimizar los errores debido a la variación del muestreo.

Las tres hipótesis usadas en cada paso de la trayectoria son: en el primer paso, los estudiantes aprenderán a reconocer y manejar la variación de muestreo utilizando datos

categoricos y estimaciones de proporciones poblacionales, y descubrirán cómo las muestras más grandes proporcionan mejores representaciones de la población. El segundo paso implica que los estudiantes conceptualicen las distribuciones de frecuencia como herramientas para explorar variaciones y predecir eventos en muestreos repetidos. asimismo, en el tercer paso, se espera que los estudiantes utilicen simulaciones para profundizar su comprensión sobre la variabilidad y mejorar la precisión de sus inferencias estadísticas sobre la población.

La construcción de la distribución muestral a través del muestreo

Silvestre et al. (2022) desarrollaron una trayectoria de aprendizaje enfocada en el razonamiento estadístico en estudiantes de bachillerato, centrándose en el muestreo y las distribuciones muestrales empíricas. El estudio hizo énfasis en conceptos clave como población, muestra, muestreo repetido y distribución muestral, y se delineó una versión preliminar del marco de desarrollo usado en esta investigación.

Diseñaron tres tareas específicas centradas en la estimación de proporciones poblacionales. Los estudiantes realizaron estas tareas utilizando simulaciones de muestreo, las cuales realizaron tanto manualmente (mediante técnicas de muestreo físico) como en software estadístico, a través de un software estadístico. Esta combinación de métodos tenía como fin proporcionar a los estudiantes una experiencia concreta, así como un entendimiento conceptual reforzado por la visualización y manipulación de datos.

Tabla 2.5

Marco de desarrollo previo de las nociones de muestreo.

Conceptos	Niveles de razonamiento		
	Ingenuo (Aditivo)	Transicional	Multiplicativa y Distribucional
Población	Material	Simulada	Matemática
Muestra	Real	Simulada	Real/Abstracta
Muestreo	Sin reemplazo	Con reemplazo	Con reemplazo
DME	Descriptiva	Frecuencial/ Probabilística	Probabilística/Estándar de variación

Los investigadores reportan que los estudiantes comenzaban con una comprensión aditiva de la muestra, es decir, concebían el muestreo repetido simplemente como un medio para aumentar el tamaño de una muestra única con la esperanza de alcanzar una inferencia más certera. Sin embargo, a medida que los estudiantes participaron en actividades que involucraban la construcción y el análisis de distribuciones muestrales empíricas, se observó una transición en su pensamiento. Empezaron a reconocer y a utilizar las propiedades estadísticas de la distribución muestral, indicativo de un cambio hacia una concepción multiplicativa. En esta etapa, los estudiantes empezaron a comprender que cada muestra es un reflejo y una reiteración independiente de un proceso de muestreo, y que las propiedades de la distribución muestral—como el centro, la dispersión y la forma—son fundamentales para comprender y hacer inferencias acerca de la población subyacente.

Esto sugiere que algunos estudiantes comenzaron a percibir la muestra no simplemente como un subconjunto de la población, sino como una entidad estadística que incorpora variabilidad y proporciona información sobre la estructura probabilística de la población desde la cual se extrae.

Capítulo III. Marco teórico

La comprensión del concepto de distribución muestral por parte de los estudiantes se da a diferentes niveles de abstracción; en este estudio se pretende documentar sobre la existencia de tres diferentes niveles. Aplicando la idea de Brandom (2000, p. 15) de que “uno no puede tener cualquier concepto a menos que uno tenga muchos conceptos” concluimos que junto con la distribución muestral conviene considerar, en un primer plano, los conceptos de población, muestra y muestreo repetido y en un segundo plano los de aleatoriedad, variación, distribución y probabilidad. Nuestra hipótesis es que los estudiantes suelen tener concepciones de tales nociones que se encuentran en algunos de los tres niveles de abstracción. Dado que nuestro objeto de estudio son los niveles de abstracción de la distribución muestral, abordaremos algunas hipótesis acerca de la abstracción y luego algunas precisiones sobre la distribución muestral. En seguida, se expondrá un marco de desarrollo de las nociones del primer plano que hemos indicado. Al final, describimos algunos aspectos relacionados con el proceso de validación del marco de desarrollo.

Marco de desarrollo cognitivo sobre concepciones

En el capítulo anterior describimos los diferentes enfoques en los que los autores clasifican las formas cada vez más sofisticadas de pensar en un tema a lo largo del tiempo. Nosotros en este estudio entendemos nuestro marco de desarrollo cognitivo como una jerarquía/trayectoria que observa tres estadios de los cuatro conceptos observados en la investigación. Es cognitivo porque observamos las formas de comprensión o concepciones de los estudiantes y profesores.

Para el concepto de "concepción" nos referimos a las creencias o teorías que los estudiantes desarrollan acerca de términos matemáticos o científicos. Según Confrey (1990), una concepción puede ser una forma inicial o ingenua de un concepto, que contiene elementos clave para una comprensión más completa en el futuro. Sin embargo, las concepciones también pueden ser incorrectas o divergentes respecto a la definición aceptada de un concepto, en cuyo caso se consideran malas concepciones o concepciones erróneas. Este estudio se enfoca en observar cómo evolucionan las concepciones de los estudiantes, buscando alinearse gradualmente con el significado normativo del concepto.

La abstracción matemática

Según Dreyfus (1991), la abstracción es un proceso clave en el desarrollo del pensamiento matemático de los estudiantes. En este proceso, se definen entidades matemáticas o conceptos de manera descontextualizada, es decir, se los separa de su contexto original o práctico para formar un sistema de ideas que se relacionan entre sí de manera formal. Estos entes abstractos, como números, figuras geométricas o funciones, se convierten en herramientas de razonamiento que no dependen directamente de sus representaciones físicas o ejemplos concretos, sino de sus propiedades y relaciones definidas dentro de la estructura matemática.

Por otro lado, Noss y Hoyles (1996) profundizan en la comprensión de la abstracción matemática al argumentar que este proceso va más allá de la simple descontextualización de un concepto. Para ellos, la abstracción es un proceso dinámico y complejo que involucra la interacción entre el conocimiento abstracto y el contexto en el que se aplica. Es decir, mientras que la abstracción implica separar un concepto de su contexto práctico inmediato, también requiere la capacidad de relacionar ese concepto abstracto de nuevo con situaciones reales o diferentes contextos. En este sentido, la abstracción no solo permite a los estudiantes comprender las estructuras y relaciones matemáticas en sí mismas, sino también aplicar ese entendimiento para interpretar y resolver problemas en diversos entornos. Para Pratt y Noss (2002), la abstracción situada es una expresión articulada de relaciones invariantes percibidas dentro de un entorno y expresadas en herramientas y formas específicas de representación.

Los conceptos de población y muestra son términos que forman parte del lenguaje ordinario y por tanto tienen referentes concretos en la vida social. Por otro lado, una distribución muestral es un ente abstracto que forma parte de un lenguaje técnico especializado que, no obstante, es un concepto que se construye a partir de las nociones de población y muestra. Pero los conceptos de población y muestra con los que se construye el concepto de distribución muestral no son los mismos que se expresan en el lenguaje ordinario, sino que tienen definiciones matemáticas propias. Así, el aprendizaje del concepto de distribución muestral puede seguir dos trayectorias, una a partir de las nociones de población y muestra del lenguaje ordinario, con sus representaciones concretas, dando lugar

a la noción de distribución muestral empírica. Esta trayectoria es la que se sigue en los estudios de Saldanha y Thompson (2007), Van Dijke-Droogers (2019) y Silvestre (2022). La otra trayectoria se propone en los libros de estadística matemática y tiene como base las definiciones formales de población y muestra, dando lugar a las distribuciones muestrales teóricas (Hogg y Craig, 1970). Pero para que este último enfoque sea asimilado por los estudiantes se puede conjeturar la existencia de un proceso que liga ambas trayectorias; una tal conjetura es el marco de desarrollo que proponemos en este trabajo.

La distribución muestral como concepto estadístico complejo

La distribución muestral es la distribución de los valores que un estadístico puede tomar en todas las posibles muestras de un tamaño específico extraídas de una población dada; los estadísticos más comunes son la media y la proporción y en este trabajo nos centraremos en la distribución muestral de proporciones. La distribución muestral representa una idea estadística compleja, ya que no se centra en una sola muestra, sino en el comportamiento de un estadístico a través de todas las muestras de un tamaño dado de una población.

Cobb (2007) destaca la naturaleza abstracta de la distribución muestral al compararla con el concepto de la función derivada en cálculo. Así como en cálculo, los estudiantes aprenden primero a calcular la derivada en un punto específico y luego avanzan hacia la comprensión de la función derivada, que describe cómo cambia la derivada a lo largo de diferentes puntos, así ocurre con la distribución muestral de un estadístico: se calcula para cada muestra y luego se avanza hacia la comprensión de la distribución del estadístico de todas las muestras. La transición para comprender la distribución muestral implica un paso en el razonamiento abstracto.

El Teorema del Límite Central es un concepto fundamental en estadística, tal como lo describe Moore (2000). Este teorema establece que, para tamaños de muestra suficientemente grandes, la distribución de los valores de ciertos estadísticos, como las medias aritméticas o proporciones, se aproxima a una distribución Normal, independientemente de la distribución de la población de origen.

Por otro lado, cuando el tamaño de la muestra no es lo suficientemente grande para cumplir con los requisitos del Teorema del Límite Central, la distribución de estos

estadísticos se aproxima mejor mediante una distribución t de Student. Esta distribución es particularmente útil para muestras pequeñas y proporciona una mejor aproximación que la distribución Normal en estos casos. Tanto la distribución Normal como la t de Student son modelos teóricos utilizados para representar la variabilidad esperada en las distribuciones muestrales. Estos modelos son herramientas esenciales en estadística para hacer inferencias sobre poblaciones a partir de muestras.

Dada una población cualquiera y un tamaño de muestra, se pueden obtener muchas muestras reales y calcular de cada una el valor de un estadístico, por ejemplo, la proporción de un atributo. La distribución de los valores así obtenidos se llama *Distribución Muestral Empírica* (DME).

Una manera para introducir en clase la noción de distribución muestral es construyendo una distribución muestral empírica con una población que se pueda manipular (Silvestre et al., 2022). Esta forma de realizar una distribución utiliza elementos físicos para simbolizar la población, como un grupo de fichas, representando así muestras como subconjuntos de esta población. Así, se forma una distribución muestral concreta a partir de estos elementos, indicando que toda población material y finita posee una distribución muestral inherente. Sin la información que ofrece el teorema del límite central, uno podría imaginar que la forma que tendría una DME dependería de la población particular en estudio y que sería impredecible; pero con la información del teorema se puede estar seguro de que si el estadístico es la proporción o la media aritmética su forma es como la distribución Normal o la t de Student.

Marco de desarrollo cognitivo para algunos conceptos de muestreo

Llamamos concepciones a las ideas que los estudiantes desarrollan sobre su mundo y a los significados propios acerca de las palabras del lenguaje y de los conceptos científicos y matemáticos que encuentran durante su aprendizaje (Confrey, 1990; Fujii, 2000). Muchas concepciones son formas que prefiguran los conceptos y que requieren más experiencias o refinamientos para consolidarse como tales. Otras concepciones mal interpretan o distorsionan los conceptos que se supone representan. Cuando los conceptos son abstractos, como los matemáticos, las concepciones que los prefiguran suelen tener referentes concretos que permiten al estudiante atribuirle significados ligados a su experiencia. En estos casos el

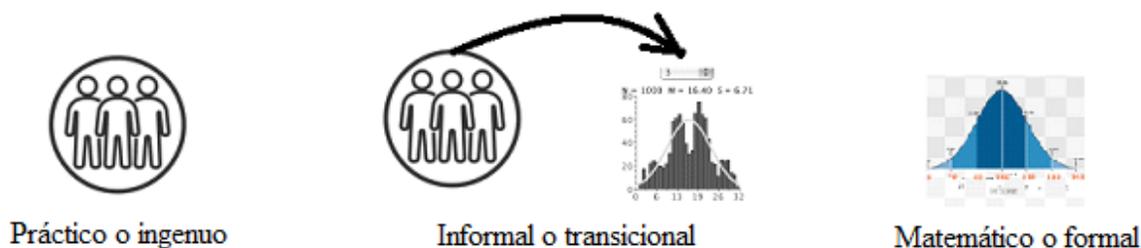
tránsito de las concepciones hacia los conceptos requiere participar en un proceso de abstracción.

Las concepciones pueden ser creencias o resultados de teorías personales de los estudiantes. El marco que proponemos se refiere a las concepciones de los estudiantes que se forman acerca de los conceptos de población, muestra, muestreo repetido y distribución muestral. Se proponen tres niveles que agrupan las diferentes concepciones de los estudiantes en torno a los conceptos elegidos; se definen de manera general de la siguiente manera (ver figura 1):

- Una concepción se clasifica en el nivel *práctico o ingenuo* cuando sus referentes son entidades concretas de la vida social del estudiante.
- Una concepción al nivel *informal o transicional* cuando incluye representaciones simbólicas ligadas a las situaciones concretas con la posibilidad de realizar operaciones matemáticas.
- Una concepción es *matemática o formal* cuando pertenece a un sistema matemático y es independiente de cualquier situación concreta.

Figura 3.1

Representación de las características de los niveles del marco



A continuación, describimos característica de los niveles de cada constructo del marco de la Tabla 3.1

Población.

Una concepción en el nivel práctico o ingenuo de población se piensa como individuos u objetos concretos o materiales; en este nivel cae la definición de Moore (2000, p. 209): “Un grupo entero de individuos sobre los que queremos obtener información”. En el nivel informal o transicional se piensa una población como representaciones numéricas o simbólicas obtenidas mediante una variable o atributo de un conjunto de individuos u objetos.

En el nivel matemático o formal se concibe a una población como una distribución teórica independiente de cualquier situación externa, pero que puede ser un modelo de una variable o atributo de un conjunto de individuos o cosas.

Tabla 3.1

Marco de desarrollo cognitivo para las nociones de muestreo

Conceptos	Nivel de respuesta		
	Práctico o Ingenuo	Informal o transicional	Matemático o formal
Población	Población material: Conjunto de personas o cosas	Valores de una variable o atributos de individuos de una población material	Una variable y su distribución
Muestra	Subconjunto de la población	Subconjunto de valores de una variable de la población	Una instancia de la n-eada de variables aleatorias X_1, X_2, \dots, X_n con $X_i \sim f(x)$, para toda i , y donde $f(x)$ es una función de densidad de probabilidad
Muestreo repetido	Obtener muestras reales sin remplazo	Obtener muestras de valores de la variable	Múltiples muestras, es decir, instancias del vector aleatorio: (X_1, X_2, \dots, X_n)
Distribución muestral (DM)	Representación de muestras reales (confusión con la distribución de la población)	Representación de valores del estadístico de muestras reales (DME)	Distribución del estadístico: Distribución muestral Teórica

Muestra

Una concepción de muestra en el nivel práctico o ingenuo es “la parte de la población que realmente examinamos con el objetivo de obtener información” (Moore, 2000, p. 209); esta caracterización resulta en una muestra sin remplazo. En el nivel informal, es un subconjunto de los valores de una variable o atributo; en esta concepción el dominio de la variable son individuos de una población concreta. En el nivel matemático o formal una muestra se define como un vector (x_1, x_2, \dots, x_n) que representan una instancia de las variables aleatorias X_1, X_2, \dots, X_n con $X_i \sim f(x)$, para toda i , y donde $f(x)$ es una función de

densidad de probabilidad (Hogg y Craig, (1970). En este caso, la muestra es con remplazo. (Según Sóbol (1976) cada x_i de (x_1, x_2, \dots, x_n) es un “sorteo” de la variable aleatoria).

Muestreo repetido

En el nivel práctico, el muestreo repetido se concibe como un proceso de obtener muestras reales sin remplazo (a nivel de individuos y de muestras) de una población material y finita, de esta manera el agregado de las muestras resulta en una muestra grande (Saldanha y Thompson, 2007; Silvestre, et al. 2022). En el nivel informal el muestreo repetido consiste en obtener muestras de valores de una variable o atributo de los individuos u objetos de muestras reales, pero el muestreo es con remplazo a nivel de muestras. En el nivel abstracto, el muestreo repetido consiste en obtener múltiples muestras (x_1, x_2, \dots, x_n) del vector aleatorio $Z = (X_1, X_2, \dots, X_n)$, con $X_i \sim f(x)$, para toda i , y donde $f(x)$ es una función de densidad de probabilidad.

Distribución muestral.

En el nivel ingenuo una distribución muestral es una representación de los valores de un estadístico de muchas muestras reales de una población de manera que contiene información de la población; con esta concepción se suele creer erróneamente que la distribución muestral se aproxima a la distribución de la población. En el nivel informal la distribución muestral es la distribución de los valores del estadístico de todas o muchas muestras de la población, la llamaremos *Distribución Muestral Empírica* (DME). En el nivel formal una distribución muestral es una distribución teórica de probabilidad que modela el comportamiento de la variación muestral y es independiente de factores externos. Cuando el estadístico es la proporción de un atributo la Distribución muestral sólo depende de la hipótesis acerca de la proporción del atributo en la población y del tamaño de la muestra.

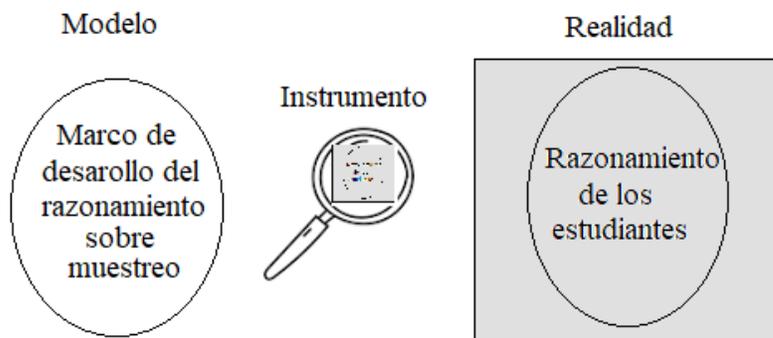
Aunque los datos de los cuales emergió la primera versión del marco (Eleazar et al., 2022) cuentan como evidencia para mostrar su plausibilidad, esta debe reforzarse con los resultados de un estudio realizado deliberadamente para validarlo. El presente estudio tiene el objetivo contribuir a dicha validación.

Sobre la validación de un marco de desarrollo

En el contexto de nuestra investigación, el objetivo principal es validar el marco de desarrollo que se refiere al razonamiento sobre conceptos de muestreo. Este marco es considerado válido si describe correctamente los estados y procesos que ocurren durante el aprendizaje de los estudiantes. Para lograr esta validación, se ha creado un instrumento específico que evalúa y explora el razonamiento sobre el muestreo. La figura 3.2 representa el modelo teórico representado por el marco, mientras que el "Razonamiento de los estudiantes" en la realidad es lo que se busca comprender y evaluar. El instrumento actúa como una herramienta que permite observar y medir cómo el razonamiento de los estudiantes se alinea con el modelo teórico. Este proceso de validación no es estándar y se evalúa en grados o niveles de acuerdo con los principios metodológicos y de interpretación, destacando que la validez de la investigación es gradual y depende de cómo se abordan y gestionan los diferentes aspectos del estudio.

Figura 3.2

Relación entre el modelo, el instrumento de evaluación y el razonamiento de los estudiantes



Dimensiones de validación

En esta investigación, proponemos un enfoque amplio para la validación, considerando múltiples dimensiones que refuerzan la solidez y calidad del estudio en cada una de sus etapas. Estas dimensiones abarcan desde los aspectos técnicos hasta la pertinencia y relevancia práctica, pasando por la validez y confiabilidad, la triangulación metodológica, el análisis fenomenológico-etiológico y el pensamiento abductivo. Cada una de estas instancias aportan a que nuestra investigación no solo sea rigurosa en su metodología, sino también relevante y aplicable.

Dimensión de los criterios técnicos

En esta dimensión partimos de los criterios descritos según Bakker (2018) que incluyen:

- Anclaje disciplinario, enfocándose en una disciplina concreta;
- Precisión en el uso de conceptos y alineación con literatura actual;
- Funcionalidad metodológica, con métodos que efectivamente respondan a la pregunta de investigación;
- Consistencia en todos los componentes del estudio; y
- Claridad en la presentación, incluyendo escritura transparente y análisis de datos sistemático.

Validez y confiabilidad

De manera general, la validez se presenta cuando uno mide lo que realmente pretende medir (Freudenthal, 1982). Sin embargo, lograr la validez en una investigación puede ser complicado, especialmente al diseñar cuestionarios eficaces. La clave de la validación recae en identificar con precisión qué se desea medir. En nuestro estudio, el foco se sitúa en los niveles de abstracción, definiendo así nuestro objeto de estudio como el 'nivel de abstracción' o 'nivel de desarrollo'. Este enfoque nos permite abordar de manera específica el tipo procesos cognitivos de los estudiantes al trabajar con los diferentes conceptos del marco.

Por otro lado, la fiabilidad (confiabilidad) se expresa cuando esa medición es replicable, es decir, ante circunstancias similares se obtienen mediciones similares (Freudenthal, 1982, Cohen et al, 2007). Las hemos incluido como parte de la misma dimensión de acuerdo a los criterios establecidos por Bakker (2018) en cada fase de la investigación. Por ejemplo, los constructos teóricos deben estar bien definidos y respaldados por la literatura. El diseño del estudio debe ser apropiado para la pregunta de investigación y replicable. La muestra debe ser representativa para garantizar la validez externa, y los procedimientos deben ser claros y repetibles para mantener la confiabilidad. Los instrumentos deben ser válidos y confiables. La recogida de datos de alta calidad, a través de grabaciones y transcripciones, asegura la validez interna y la confiabilidad. La triangulación de datos aumenta la validez interna, mientras que la transparencia en la argumentación y la coherencia en las conclusiones garantizan la validez inferencial y la confiabilidad del estudio (ver tabla 3.2).

Tabla 3.2*Validez y confiabilidad en las fases de la investigación (Adaptado de Bakker, 2018, p.90)*

Aspecto del estudio	Validez	Fiabilidad
Constructos teóricos	¿Bien definido en línea con la literatura?	¿Se pueden encontrar las fuentes?
Diseño y procedimiento de investigación	¿Es el diseño de investigación adecuado para la pregunta planteada?	¿Puede replicarse (virtualmente) el procedimiento?
Procedimiento de muestreo	¿Muestra representativa? (generalización; validez externa)	¿Puede replicarse el procedimiento de muestreo; es claro?
Instrumentos	¿Qué caso es? ¿Es un instrumento válido? (validez de constructo; validez consecuente)	¿Es un instrumento fiable?
Recolección de datos	¿Son los datos de alta calidad? (validez interna)	Hacer grabaciones de audio y/o vídeo para evitar problemas de memoria
Análisis de datos	¿Se ha aplicado la triangulación de datos o la verificación por miembros? (validez interna)	Hacer transcripciones usando un esquema de codificación y midiendo el acuerdo inter-evaluador.
Conclusión de resultados	¿Se han extraído conclusiones de una manera válida? (validez inferencial)	¿Es transparente la argumentación? ¿Llega el otro investigador a las mismas conclusiones?

Triangulación de métodos

Dentro de las dimensiones de validación consideradas, destacamos la triangulación en general como la tercera dimensión. La triangulación no se limita a un único método o perspectiva; en cambio, involucra la aplicación de una variedad de métodos, fuentes, teorías o incluso diferentes investigadores para examinar un fenómeno (Cohen et al., 2007). Esta diversidad en los enfoques y puntos de vista permite alcanzar una comprensión más profunda y detallada del tema investigado. Al integrar múltiples perspectivas, la triangulación ayuda a contrarrestar las limitaciones o sesgos que podrían surgir al utilizar un solo método o fuente.

Pertinencia

En la estructura de las dimensiones de validación, identificamos la pertinencia como la cuarta dimensión clave. Esta propuesta se basa en la importancia que Freudenthal (1982)

le otorga a la pertinencia en la investigación, destacando que los resultados y hallazgos de la investigación deben tener una aplicación directa y beneficiosa en la enseñanza real. La pertinencia, en este sentido, asegura que la investigación trascienda el ámbito académico y contribuya de manera concreta y positiva al campo de la educación, afectando tanto la teoría como la práctica educativa. La inclusión de la pertinencia en el ciclo de validación subraya la importancia de que los estudios en educación matemática no solo avancen en el conocimiento teórico, sino que también tengan un impacto práctico significativo en el aula.

De lo fenomenológico a lo etiológico

La quinta dimensión de validación, según Freudenthal (1982), implica combinar enfoques fenomenológicos y etiológicos. Un enfoque fenomenológico, se centra en observar y describir los fenómenos educativos tal como se presentan, proporcionando una comprensión básica de los problemas y situaciones en la educación. Posteriormente, se puede dar un paso hacia un análisis etiológico, que busca entender las causas y factores subyacentes detrás de los fenómenos observados. Este enfoque profundiza en las razones detrás de la efectividad o ineffectividad de métodos de enseñanza y las reacciones de los estudiantes, permitiendo así explicaciones más completas y fundamentadas en la investigación educativa.

Abducción

La sexta dimensión del ciclo de validación, inspirada en las ideas de Charles Sanders Peirce, nos centramos en la abducción como un elemento importante en el proceso de una investigación. A diferencia de la inducción, que se basa en el uso de evidencia empírica para apoyar una hipótesis, la abducción introduce hipótesis al comienzo de la indagación. Peirce ve la abducción como un proceso que invita a considerar hipótesis para su evaluación, sin comprometerse con ellas desde el principio (Stanley & Nyrup, 2020). Este enfoque se orienta a identificar aquellas hipótesis que merecen ser exploradas más profundamente, independientemente de su plausibilidad o probabilidad inicial. Incluso, en ocasiones, una hipótesis improbable puede ser valorada y aceptada provisionalmente para su posterior verificación y análisis.

Relaciones entre dimensiones

En las dimensiones de validación propuestas, los criterios técnicos establecen una base sólida y coherente, asegurando que todos los aspectos de la investigación, desde la formulación de la pregunta hasta la metodología y el análisis de datos, sean rigurosos y relevantes. Esta base técnica favorece los aspectos de validez y confiabilidad de la segunda dimensión. La triangulación, se apoya en esta solidez metodológica para explorar el fenómeno de estudio de manera más completa, ofreciendo así una visión más rica y matizada.

La pertinencia conecta el estudio con su aplicación práctica, asegurando que los hallazgos sean útiles y relevantes para la comunidad o el campo de estudio. Esta relevancia práctica se ve reforzada por el enfoque fenomenológico-etiológico, que comienza con una visión detallada de los fenómenos tal como se presentan y luego profundiza en sus causas subyacentes, proporcionando una comprensión más profunda y aplicable.

Por último, la abducción introduce nuevas hipótesis y direcciones para la investigación. Este proceso de generación de hipótesis nutre las dimensiones creando un ciclo de validación, permitiendo que el estudio se adapte y evolucione en respuesta a nuevos descubrimientos y perspectivas. En conjunto, estas seis dimensiones trabajan de manera sinérgica para crear un proceso de validación que asegure la calidad y relevancia del estudio.

Capítulo IV. Método

El propósito original de la presente investigación es generar evidencia para apoyar la validez del marco de desarrollo para las nociones de muestreo formulado en el capítulo anterior. Decimos que dicho marco tiene validez si prevé estados y procesos que ocurren efectivamente como resultado de, o durante, el proceso de aprendizaje de los temas de muestreo por parte de los estudiantes. Evitamos considerar que la validez de este marco de desarrollo se pueda probar de forma absoluta, más bien creemos que es necesario proceder por aproximaciones sucesivas y de manera abductiva creando evidencias desde un enfoque y otro para hacer de su acumulación un argumento convincente.

Un problema implicado es la elaboración de instrumentos para obtener datos pertinentes que apoyen o no las proposiciones del marco. Con este objetivo, se diseñó un cuestionario que busca explorar el razonamiento de los estudiantes acerca de los cuatro conceptos clave del marco: población, muestra, muestreo repetido y distribución muestral. Este cuestionario incluye preguntas junto con posibles respuestas que desde un análisis a priori representan a los diferentes niveles de razonamiento del marco. Algunas de las preguntas del cuestionario son adaptaciones de preguntas formuladas en trabajos previos relacionados con distribuciones muestrales, como los estudios realizados por Chance et al. (2004) y Batanero et al. (2018).

Ahora bien, es necesario plantear un problema previo ¿El instrumento que hemos elaborado es válido? El propósito del cuestionario es permitir reunir datos que informen sobre el marco de desarrollo y entonces es válido si cumple con dicho propósito. En consecuencia, el problema original nos llevó a este problema previo que es al que nos concentraremos en este estudio, asumiendo que su respuesta contribuye a apoyar el problema original. A partir de este momento nos referiremos a la validez del instrumento la mayoría del tiempo, salvo excepciones y en una reflexión final regresaremos al asunto de la validez del marco.

Para el proceso de validación del instrumento, se llevaron a cabo cinco fases específicas con el objetivo de recopilar datos:

1. Entrevista a una profesora experta en estadística.
2. Primera aplicación del cuestionario.

3. Valoración del cuestionario por especialistas en educación estadística.
4. Aplicación del cuestionario a dos modelos de lenguaje basados en Inteligencia Artificial.
5. Segunda aplicación del cuestionario, en su versión mejorada, a profesores de CCH.

Es importante mencionar que el cuestionario, utilizado como instrumento de recolección de datos, fue actualizado, y presumiblemente mejorado, en cada etapa según las necesidades identificadas (consultar Apéndice F).

Incluimos como parte de la validación las evidencias que permitan argumentar que el marco y el instrumento ofrecen información útil para ampliar la investigación sobre el razonamiento estadístico de los estudiantes y profesores.

En las secciones subsiguientes, describimos los elementos metodológicos concretos tales como los sujetos de estudio, las herramientas empleadas, la metodología adoptada, las pruebas recabadas y la manera de interpretación de la información asociada a cada una de las cinco fases realizadas.

Entrevista a una profesora especializada en estadística

Tras analizar el marco de desarrollo derivado de investigaciones anteriores, se elaboró un cuestionario destinado a actuar como herramienta para la recopilación de datos en el proceso de validación. El 31 de mayo de 2022, se llevó a cabo una entrevista con una investigadora y profesora en el campo de la estadística, utilizando la plataforma de videoconferencias Zoom.

El propósito principal de esta entrevista era obtener una perspectiva experta sobre el marco de desarrollo y el instrumento diseñado. Esta interacción no solo ofreció una oportunidad para adquirir una visión desde el ámbito académico de la profesora, sino también permitió identificar omisiones o aspectos técnicos que requerían revisión. Además, la conversación facilitó la clarificación de aspectos específicos relacionados con las preguntas del cuestionario y permitió contrastar nuestras interpretaciones con las de la experta. A partir de la retroalimentación proporcionada por la doctora, se realizaron modificaciones en el diseño de ciertos ítems del cuestionario.

Sujeto

La profesora entrevistada es catedrática del departamento de matemáticas de la facultad de ciencias en una universidad pública en México. Ella imparte cursos de estadística de forma regular a varios niveles y tiene más de 10 años de experiencia. Su participación en la investigación se debió a su conocimiento y experiencia en la materia objeto de estudio, además de la disposición e interés para colaborar en la presente investigación. Esta profesora fue la única que respondió positivamente a una solicitud que el autor de esta tesis hizo a varios profesores experimentados de estadística para participar en una entrevista para preguntarles sobre el marco de la presente investigación.

Herramientas

Se empleó una versión inicial del instrumento de recolección de datos, que se encuentra detallada en el Apéndice A. Este cuestionario consta de cinco preguntas de opción múltiple. El primer ítem aborda el concepto de población, el segundo se centra en el concepto de muestra, el tercer ítem se relaciona con el muestreo repetido, mientras que los ítems 4 y 5 se enfocan en el concepto de distribución muestral.

Tabla 4.1

Redacción del ítem 1 del cuestionario presentado a la experta en estadística

1. ¿Qué ejemplo coincide mejor con el concepto que tienes de población estadística?

- a) Los estudiantes de la facultad.
 - b) Todos los posibles resultados de lanzar una moneda 10 veces.
 - c) Una caja con 500 pelotas negras y 800 blancas.
 - d) Una distribución generada por una variable aleatoria.
 - e) Otra. Explica tu respuesta
-

En la tabla 4.1 aparece el ítem 1 del cuestionario usado en esta primera fase, el cual indaga sobre el nivel de razonamiento del concepto de población, y cada opción representa diferentes niveles de abstracción de acuerdo con el marco establecido. La opción a) refleja una interpretación práctica o empírica del concepto, mientras que la b) lo vincula de manera más abstracta con el espacio muestral de un experimento. La opción c) proporciona un ejemplo numérico de una población específica que igualmente se encontraría en un nivel práctico, y la d) representa la interpretación más abstracta conforme al marco. Además, se incluye una última opción abierta para aquellos que posean una perspectiva diferente.

Procedimiento

La entrevista consistió en una videollamada en la cual se llevó a cabo la presentación del marco de desarrollo y la explicación del objetivo del cuestionario; después la experta expresó su opinión sobre algunas preguntas del cuestionario y opinaba si las consideraba claras y pertinentes.

Fuente de datos

El audio de la entrevista de casi dos horas se grabó y posteriormente se transcribió para su posterior análisis.

Análisis de los datos

Se empleó una tabla para dividir la entrevista por secciones y se discutieron los principales tópicos y observaciones realizados por la investigadora. Mediante el análisis se destacaron los siguientes aspectos: explicaciones sobre conceptos estadísticos y su enseñanza, conceptos ausentes en el marco, críticas al marco y afirmaciones controversiales.

Primera exploración con el instrumento de toma de datos

Tras la entrevista con la experta en estadística, procedimos a realizar una aplicación exploratoria inicial del cuestionario con estudiantes entre el 20 y el 29 de julio de 2022. El objetivo principal de esta fase exploratoria era evaluar la capacidad del cuestionario para capturar información sobre el razonamiento de los estudiantes. Esta exploración nos permitió identificar áreas de mejora tanto en el instrumento como en el marco conceptual subyacente.

Sujetos

La muestra utilizada en este estudio estuvo conformada por cuatro participantes, de los cuales uno era mujer y tres eran hombres. Los participantes tenían edades de 22, 22, 23 y 23 años, respectivamente. En cuanto a su formación académica, dos de ellos estudiaban matemáticas aplicadas, mientras que los otros pertenecían a la carrera de matemáticas. Los criterios de inclusión fueron: ser estudiante y haber llevado un curso de estadística. No se establecieron criterios de exclusión. La convocatoria y la aplicación del cuestionario se hicieron vía Internet porque todavía había restricciones para realizar clases presenciales por motivos de la pandemia. Como la aceptación a la convocatoria era voluntaria sólo se logró obtener respuestas de cuatro estudiantes.

Herramientas

Para esta fase, se implementó un cuestionario compuesto por nueve preguntas, las cuales se encuentran detalladas en el Apéndice B. En este cuestionario, se incorporaron preguntas de respuesta abierta relacionadas con cada uno de los conceptos del marco de desarrollo, siguiendo las recomendaciones que hacen Cohen et al. (2007). Además, posterior a cada pregunta de opción múltiple, se incluyó una pregunta complementaria que instaba a los estudiantes a explicar las razones detrás de su elección, con el objetivo de obtener respuestas más detalladas. En la tabla 4.2, se presenta un extracto que incluye los ítems relacionados con el concepto de muestra.

Tabla 4.2

Ítems 3 y 4 del cuestionario de contenido aplicado a estudiantes

3- ¿Qué entiendes por una muestra en el contexto de la estadística?

(respuesta abierta)

4- ¿Qué ejemplo coincide mejor con el concepto que tienes de muestra estadística? (Puedes elegir más de una opción)

- a) Todos los casos de un lanzamiento de dos dados.
- b) Un conjunto de n valores de una variable aleatoria elegidos al azar de acuerdo con su distribución de probabilidad.
- c) Un grupo de n alumnos elegidos al azar de una preparatoria con M alumnos.
- d) Cualquier subconjunto de individuos de la población.
- e) Otra. Explica tu respuesta abajo.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior.

(Respuesta abierta)

La Tabla 4.2 muestra los ítems 3 y 4 del cuestionario dirigido a los estudiantes. En la pregunta 3, se busca obtener una visión general de la concepción del estudiante sin sugerir ningún punto específico. Por otro lado, en la pregunta 4, se intenta identificar los diferentes niveles de razonamiento. El inciso a) presenta una opción relacionada con el espacio muestral, que podría detectar a los estudiantes que no distinguen entre muestra y población en situaciones de azar. En el inciso b), se formula una versión matemática de muestra, que es la que se utiliza en los procesos de simulación computacional de muestreo repetido. En el inciso c), se adopta una perspectiva intermedia entre las versiones concreta y abstracta de muestra, que considera la aleatoriedad y la probabilidad de selección de cada individuo. Finalmente, la opción d) ofrece una perspectiva acorde a una versión práctica del concepto en cuestión.

Procedimiento

La aplicación del cuestionario piloto se realizó a través del envío de un formulario de Google. Se les proporcionó una breve explicación sobre el objetivo de la investigación y se les pidió que respondieran el cuestionario de manera sincera, asegurándoles la confidencialidad de sus respuestas. Se permitió un plazo de dos a tres días para que los estudiantes lo respondieran.

Fuente de datos

La fuente de datos son las respuestas que proporcionaron los estudiantes ante las preguntas del cuestionario, mismas que fueron vaciadas en un editor de texto para poder llevar a cabo el análisis.

Análisis de los datos

Para el análisis de los datos recopilados a través de la primera exploración, se llevó a cabo una revisión de las respuestas de los participantes por ítem del cuestionario, con el fin de identificar patrones y categorías que pudieran estar relacionadas con los conceptos del marco. A partir de esta revisión, se procedió a categorizar las respuestas en función de los diferentes niveles del marco de desarrollo

Valoración de expertos en educación estadística

Tras la fase exploratoria del cuestionario, procedimos a someter el instrumento a la evaluación de expertos en educación estadística. Esta valoración se desarrolló entre diciembre de 2022 y enero de 2023. El propósito principal de consultar a estos expertos era fortalecer la validez del cuestionario, asegurando su relevancia y adecuación en el ámbito de la educación estadística.

Sujetos

Los tres expertos en educación estadística invitados para revisar el marco y el cuestionario son investigadores en educación estadística con un amplio currículo. Una de ellas tiene una amplia trayectoria de más de 30 años en la investigación en educación estadística, otro con más de 10 años y el tercero con más de 5 años.

Herramientas

Para esta etapa se elaboró un documento (Apéndice D) que incluía la exposición del marco de desarrollo y para cada pregunta del instrumento se les pedía a los expertos que llenaran una tabla para valorar cada uno de los ítems en tres ámbitos: formulación, pertinencia y relevancia. Además, se añadió un espacio para dejar comentarios (Ver tabla 4.3).

Durante la elaboración del cuestionario dirigido a los evaluadores, se tuvo en cuenta la referencia de ejemplos proporcionados por Batanero y Hernández (Comunicación personal) y Malaspina (Comunicación personal), quienes han aplicado cuestionarios para obtener información relevante de expertos en educación estadística.

Tabla 4.3

Tabla para valoración de los ítems sobre muestra por parte de los expertos

No.	CONCEPTOS / ÍTEMS	Formulación			Pertinencia			Relevancia			Observaciones / Sugerencias
		Adecuada	A mejorar	No Adecuada	Pertinente	Con dudas	No Pertinente	Relevante	Con dudas	No Relevante	
	Concepto: Muestra										
4	¿Qué entiendes por una muestra en estadística?										
5	En el mismo contexto del problema 1 sobre los tiros libres en el basquetbol; se quiere obtener una muestra de tamaño 5. ¿Cuál de los siguientes incisos elegirías como muestra?										
6	Se quiere obtener una muestra de tamaño 10 de una urna con 500 fichas blancas y 800 fichas negras. ¿Cuál de los siguientes incisos elegirías como muestra?										

Procedimiento de aplicación

Se envió el formato para la evaluación vía correo electrónico y se les dio un mes para que lo devolvieran cumplimentado y con comentarios.

Fuente de datos

Una vez obtenidas las respuestas de los expertos, se procedió a elaborar un cuadro comparativo de las valoraciones a cada ítem del cuestionario. Este cuadro incluía las respuestas individuales de cada experto para cada una de las evaluaciones de los ítems.

Análisis de los datos

Las respuestas a cada valoración de los ítems fueron clasificadas en tres categorías: "no propone modificaciones", "aportó algún comentario" y "desaprueba el ítem". Se visualizaron de forma global en un cuadro comparativo. Con esto pudimos identificar ítems que podían estar sujetos a revisión.

Aplicación del cuestionario a dos Modelos de Lenguaje con Inteligencia Artificial

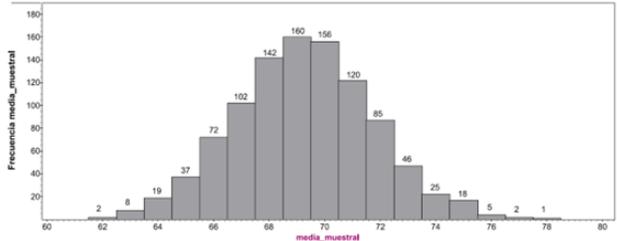
Tras recibir la retroalimentación de los expertos, avanzamos en la implementación del cuestionario a dos modelos de lenguaje basados en inteligencia artificial. Esta aplicación se llevó a cabo entre el 6 y el 12 de julio de 2023. El objetivo principal de esta fase era verificar la capacidad del instrumento para discernir entre dos niveles distintos de razonamiento. Esta diferenciación es esencial para aportar a la validez del cuestionario.

Sujeto

Hemos recurrido al análisis de las respuestas de dos Modelos de Lenguaje Grandes (LLM por sus siglas en inglés): el GPT-3.5 y el GPT-4. Los modelos de lenguaje GPT son instancias de la Inteligencia Artificial con la capacidad de procesar y generar lenguaje humano, ofreciendo una interacción similar a la que tendría lugar entre humanos que conversan a través de un chat como WhatsApp o Messenger. Estos modelos se encuentran disponibles en la página web de ChatGPT de la empresa OpenAI. Es importante precisar que la versión GPT-4 es de pago.

Tabla 4.4

Ejemplo de descripción verbal de la gráfica del ítem 12 para usarla en los modelos de lenguaje

Distribución muestral simulada	Verbalización de la distribución																																				
 <table border="1"><caption>Datos del histograma</caption><thead><tr><th>media_muestral</th><th>Frecuencia</th></tr></thead><tbody><tr><td>62</td><td>2</td></tr><tr><td>63</td><td>8</td></tr><tr><td>64</td><td>19</td></tr><tr><td>65</td><td>37</td></tr><tr><td>66</td><td>72</td></tr><tr><td>67</td><td>102</td></tr><tr><td>68</td><td>142</td></tr><tr><td>69</td><td>160</td></tr><tr><td>70</td><td>156</td></tr><tr><td>71</td><td>120</td></tr><tr><td>72</td><td>85</td></tr><tr><td>73</td><td>46</td></tr><tr><td>74</td><td>25</td></tr><tr><td>75</td><td>18</td></tr><tr><td>76</td><td>5</td></tr><tr><td>77</td><td>2</td></tr><tr><td>78</td><td>1</td></tr></tbody></table>	media_muestral	Frecuencia	62	2	63	8	64	19	65	37	66	72	67	102	68	142	69	160	70	156	71	120	72	85	73	46	74	25	75	18	76	5	77	2	78	1	$\{(Media\ muestral, Frecuencia): (62,2), (63,8), (64,19), (65,37), (66,72), (67,102), (68,142), (69,160), (70,156), (71,120), (72,85), (73,46), (74,25), (75,18), (76,5), (77,2), (78,1)\}$
media_muestral	Frecuencia																																				
62	2																																				
63	8																																				
64	19																																				
65	37																																				
66	72																																				
67	102																																				
68	142																																				
69	160																																				
70	156																																				
71	120																																				
72	85																																				
73	46																																				
74	25																																				
75	18																																				
76	5																																				
77	2																																				
78	1																																				

Herramientas

El cuestionario empleado incorpora las correcciones sugeridas por los expertos. Dado que los modelos no pueden procesar imágenes, se verbalizaron los elementos gráficos presentes en los ítems 11 y 12, como las distribuciones (ver tabla 4.4). Esto permitió que los modelos pudieran seleccionar una opción basada en el texto. Además, se incluyó una pregunta abierta como ítem 10 para explorar la concepción sobre distribución muestral.

Tabla 4.5

Ejemplo de diálogo con el modelo GPT-4 en la página web, centrado en el ítem 1.

🔗 GPT-4

PE

¿Qué entiendes por una población en estadística?

En estadística, una población se refiere a la totalidad de individuos, objetos, eventos o medidas que tienen al menos una característica en común y que son de interés para un estudio o una investigación. Esta población puede variar enormemente en tamaño y naturaleza, dependiendo del contexto.

Por ejemplo, si un investigador está interesado en estudiar la altura promedio de los estudiantes universitarios en España, entonces la población sería todos los estudiantes universitarios en España. Si alguien quiere investigar el número de accidentes de tráfico en Nueva York en el último año, la población sería todos los accidentes de tráfico ocurridos en Nueva York durante ese período.

Además, hay que tener en cuenta que en estadística se utilizan muestras, que son subconjuntos de la población, para realizar inferencias sobre la población total. Esto se hace debido a que a menudo es impracticable o imposible investigar a cada individuo u objeto en la población.

🔄 Regenerate

Procedimiento

Utilizamos la página web de ChatGPT (<https://chat.openai.com/>) para acceder a los dos modelos de lenguaje empleados en esta fase de la investigación. Una vez en la interfaz, seleccionamos el modelo específico al que deseábamos hacer las preguntas. Posteriormente, escribimos la redacción de cada uno de los 12 ítems del cuestionario en el cuadro de diálogo de ChatGPT (ver tabla 4.5), con el propósito de obtener las respuestas de los modelos a cada uno de los ítems.

Fuente de datos

Cada una de las respuestas recibidas por los modelos de lenguaje fue registrada en una tabla de un editor de texto para su posterior clasificación y análisis. Es decir, para cada uno de los modelos se tiene una tabla con su respuesta al ítem y la valoración con respecto al marco de desarrollo por parte de nosotros y una serie de comentarios.

Análisis de los datos

Para el análisis de los datos lo que hicimos fue realizar una tabla comparativa como en los casos anteriores y encontrar los patrones y cuestiones importantes relacionadas con nuestro objetivo de validación.

Segunda exploración con profesores de CCH

Posterior a la evaluación por expertos y de forma paralela a la aplicación del instrumento a los modelos de lenguaje, se realizó la aplicación del instrumento a profesores del Colegio de Ciencias y Humanidades (CCH) de la UNAM. La aplicación se realizó entre el 9 y 24 de julio de 2023. Tuvo como propósito ver la forma en que el instrumento podría captar la variabilidad y el nivel de razonamiento de los profesores. Los datos permitieron hacer descubrimientos sobre ciertas relaciones entre los conceptos del marco.

Sujetos

Once docentes de los cinco planteles del CCH participaron en un curso intersemestral sobre tecnología en pruebas de significación, donde se les aplicó el cuestionario. Sus formaciones académicas abarcan áreas como física, actuaría, matemáticas, ingeniería y química. Cinco poseen maestría, tres doctorado y tres solo licenciatura. La mayoría ha enseñado cursos de estadística y sus edades oscilan entre 30 y 60 años.

Herramientas

La aplicación del cuestionario a profesores de CCH (ver apéndice E) se llevó a cabo a través del envío de un formulario de *Google Forms*. El cuestionario está compuesto por 12 ítems, que se dividen en los cuatro grupos. La primera pregunta de cada grupo es abierta y se enfoca en la concepción que los profesores tienen sobre los conceptos del marco. Las dos preguntas siguientes en cada grupo son de opción múltiple con un complemento de pregunta en donde se pide una explicación abierta.

Tabla 4.5

Ejemplo de la redacción del ítem 8 usado el cuestionario con profesores de CCH.

8.- Se quiere realizar un muestreo repetido de tamaño 100 en el contexto de los tiros a la canasta de basquetbol. ¿Cuál de las siguientes opciones elegirías para hacerlo?

- a) Tomar a un alumno al azar de la escuela y registrar el número de anotaciones en 10 tiros. El alumno elegido ya no puede ser seleccionado nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- b) Tomar a un alumno al azar de la escuela, se registra el número de anotaciones en 10 tiros. El alumno elegido podría ser seleccionado nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- c) Tomar 10 alumnos al azar de la escuela, se registra el número de anotaciones en 10 tiros de cada uno. Los 10 alumnos elegidos no pueden ser seleccionados nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- d) De la lista de todos los alumnos se eligen 10 alumnos al azar de la escuela, se registra el número de anotaciones en 10 tiros de cada uno. Se toma otra muestra de la misma manera (algunos podrían volver a caer en la muestra). Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- e) Otro procedimiento. Explica tu respuesta abajo.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior (Respuesta abierta)

La tabla 4.5 muestra la redacción del ítem 8, que es una pregunta de opción múltiple relacionada con el concepto de muestreo repetido. Las respuestas incluyen dos opciones de muestreo sin reemplazo y dos con reemplazo. La opción b) es la más abstracta, ya que representa un muestreo con reemplazo individual. La opción a) y c) ejemplifican muestreos sin reemplazo. Mientras que en la opción d) aunque los bloques de 10 alumnos son independientes entre sí, dentro de cada bloque hay una dinámica de muestreo sin reemplazo, lo que afecta la probabilidad de selección a medida que se avanza dentro del bloque.

Procedimiento

Se envió el cuestionario diseñado en la plataforma Google Forms a su correo y se dio dos semanas de plazo para responder el cuestionario.

Fuente de datos

Se vaciaron las respuestas de los once a profesores a las doce preguntas del cuestionario en una tabla en el editor de texto para su análisis.

Análisis de los datos

Con los datos de las respuestas por ítem, se clasificaron de acuerdo con las categorías preestablecidas de nivel de razonamiento. A continuación, se elaboró una tabla comparativa en la que cada fila representaba el nivel de respuesta, conforme a la clasificación, para cada ítem evaluado por los profesores. Para profundizar en el análisis, se trianguló la información, organizando las columnas de los ítems desde aquellos con mayor frecuencia de respuestas prácticas hasta los de menor frecuencia. Finalmente, se identificaron grupos específicos dentro de la tabla y se procedió a categorizar dichos grupos de manera sistemática.

Capítulo V. Presentación de resultados

Dentro de este capítulo, se muestran los hallazgos derivados de las cinco etapas de recolección de datos. Estas fases incluyen: una entrevista dirigida a una académica universitaria con especialización en estadística; la aplicación del instrumento a un conjunto de cuatro estudiantes que han cursado estadística; la valoración y perspectiva de tres expertos en la investigación educación estadística; la experimentación con los modelos de lenguaje GPT-3.5 y GPT-4 mediante el citado instrumento; y finalmente, la aplicación del cuestionario a un grupo de once docentes en activo del ámbito de nivel medio superior. A continuación, se detalla cada uno de los resultados de las acciones realizadas.

Entrevista a una Especialista en Estadística

Se solicitó a la profesora, experta en estadística, que ofreciera su perspectiva y evaluara el marco de desarrollo sobre nociones de muestreo. Además, se le pidió que analizara el cuestionario diseñado, cuyo objetivo principal es detectar los niveles de razonamiento.

La entrevista de aproximadamente dos horas se transcribió y analizó. Mediante el análisis se destacaron los siguientes aspectos:

1. Explicaciones sobre conceptos estadísticos y su enseñanza
2. Conceptos ausentes en el marco
3. Críticas al marco
4. Afirmaciones controversiales

A continuación, procederemos a desglosar cada uno de los cuatro puntos destacados durante el análisis de la entrevista.

Explicaciones sobre conceptos estadísticos y su enseñanza.

A lo largo de la entrevista la profesora consideró pertinente ofrecer su punto de vista sobre los rasgos característicos de los conceptos de población, muestra, muestreo repetido e inferencia. Durante la conversación, la experta subrayó la relevancia de definir adecuadamente la población en cualquier estudio estadístico. Argumentó que esto es esencial

porque la meta principal de la inferencia estadística es proyectar o generalizar los hallazgos obtenidos de una muestra específica a la población completa de la que se extrajo.

Al hablar sobre el concepto de muestra, la experta hizo hincapié en determinar con precisión la probabilidad con la que cada individuo es seleccionado para formar parte de la muestra, asegurando así una representatividad adecuada.

Además, hizo una distinción clara entre el proceso de inferencia y el muestreo repetido. Aunque ambos son importantes en estadística, recaló que una inferencia se realiza con base en una única muestra y que el muestreo repetido no es parte inherente del proceso de la inferencia. Como el análisis de la muestra no es en la perspectiva de llevar a cabo simulaciones para generar distribuciones muestrales, sino más bien seguir el procedimiento canónico utilizando en la tabla de la distribución normal, considera al concepto de muestreo repetido como ajeno a dicho procedimiento.

Conceptos ausentes en el marco

Durante la revisión del marco, la profesora destacó la omisión de conceptos clave como la aleatoriedad, el marco muestral y las técnicas de muestreo, resaltando la importancia de incluirlos. Enfatizó que la comprensión del muestreo requiere primero entender la aleatoriedad, ya que, sin ella, la inferencia estadística pierde su validez. Además, subrayó la necesidad de considerar la probabilidad de selección de cada individuo para realizar estimaciones precisas. También resaltó la diversidad de técnicas de muestreo utilizadas profesionalmente, como el muestreo estratificado en encuestas electorales, enfatizando su relevancia tanto en el ámbito académico como en el profesional.

Críticas al marco

La profesora expresó discrepancias con ciertas definiciones en el marco de desarrollo, criticando la definición de "población" como mera distribución matemática. Argumentó que la noción de población en ciertos contextos se vincula con el concepto de espacio muestral, pero esta relación posee una complejidad abstracta que podría confundir a los estudiantes novatos. Además, cuestionó la viabilidad del muestreo repetido para la inferencia, dado su alto costo y complejidad, señalando que las muestras simuladas representan un nivel conceptual más avanzado, potencialmente confuso para los alumnos. En cuanto al muestreo

con y sin reemplazo, sostuvo que no hay diferencia de nivel entre ambos, dependiendo de la elección del investigador y del contexto de la investigación.

Afirmaciones controversiales

Se identificaron algunas declaraciones de la profesora que resultaron ser ambiguas o discutibles. Una de las afirmaciones que llamó la atención fue su postura de que “la inferencia estadística carece de sentido sin la presencia de la aleatoriedad en el fenómeno bajo estudio”. Esta perspectiva es cuestionable, ya que la estadística no solo se ocupa de fenómenos aleatorios, sino también de aquellos que son determinísticos. Es importante destacar que la aleatoriedad en la estadística a menudo proviene del proceso de muestreo y no necesariamente del fenómeno en sí.

La profesora también mencionó una relación entre la población y el espacio muestral de una variable aleatoria, pero no profundizó en su especificidad. Para que un espacio muestral pueda ser considerado como una representación de la población se debe conocer la frecuencia o probabilidad con la que se selecciona un elemento de dicho espacio. En otras palabras, es necesario el conocimiento de la distribución. Si la maestra se diera cuenta que el espacio muestral no es suficiente para caracterizar una población, sino que se requiere saber la probabilidad de cada punto muestral concluiría que contradice su anterior afirmación de que una población no es equivalente a una distribución.

La descripción anterior sobre las opiniones de la profesora revela una interpretación basada en su experiencia y conocimientos sobre los conceptos clave que un estudiante debería adquirir al abordar el tema de muestreo. Ella concibe el marco como un temario para cubrir el contenido estadístico de la inferencia y el instrumento como un examen para evaluar a los estudiantes. Su enfoque está influenciado por su trayectoria como docente en estadística, priorizando la corrección de los conceptos presentados. Mientras reconoce la importancia fundamental de los conceptos de población y muestra en la enseñanza del muestreo, manifiesta reservas sobre la relevancia de introducir conceptos como el muestreo repetido y la Distribución Muestral Simulada (DMS) en el currículo. En resumen, su postura refleja una combinación de rigor académico y cierto pragmatismo pedagógico basado en su experiencia docente.

La profesora apoyó la elaboración de instrumentos para apoyar la labor del profesor. También estuvo de acuerdo en la importancia de evaluar el razonamiento de los estudiantes sobre los conceptos de población y muestra.

Primera exploración con el instrumento

Se invitó a estudiantes universitarios, quienes previamente habían cursado una asignatura de estadística, a participar en la solución al cuestionario (Apéndice B). Dadas las circunstancias de la pandemia que aún afectaba en 2022, la encuesta se llevó a cabo de manera remota. Sin embargo, de todos los estudiantes contactados, sólo cuatro decidieron participar y completar la encuesta. En la tabla 5.2 se presentan los niveles de sus repuestas a las preguntas, Con el objetivo de preservar la confidencialidad y privacidad de los participantes, se le asignó un nombre ficticio a cada uno. Así, los estudiantes que participaron en la encuesta son referidos en este estudio como Raúl, Susana, Javier y Daniel. A continuación, una descripción individual de las respuestas de los estudiantes.

Respuestas por parte de Raúl

Tablas 5.1

Clasificación de las respuestas de Raúl

Ítems	Población			Muestra		Muestreo repetido		Distribución muestral	
	1	2	3	4	5	6	7	8	9
Raúl	P	P	P	P	P	P	P	F	O

Raúl ha proporcionado respuestas que se han clasificado de la siguiente manera: "prácticas" en los ítems del 1 al 7, "concepción formal" en el ítem 8 y, en el ítem 9, su respuesta no ha podido ser clasificada dentro del marco de desarrollo establecido. En particular, en su respuesta al ítem 2, donde se le solicitó seleccionar las opciones que mejor representaran su propio concepto de población, Raúl optó por las opciones a) "Los estudiantes de la facultad" y c) "Una caja con 500 pelotas negras y 800 blancas". Justificó su elección argumentando que estas dos opciones permiten la extracción de información relacionada con la altura, sexo y carrera en el caso de los estudiantes, así como el color de

las pelotas en el caso de la caja. Esto indica que Raúl no considera los ejemplos más abstractos, lo que lo sitúa en el nivel práctico según nuestra clasificación.

En el ítem 8, donde se planteó la pregunta sobre la naturaleza de la distribución que surgiría al tomar 500 muestras de una distribución dada. Raúl seleccionó la opción del inciso c y formuló su respuesta de la siguiente manera: "Lo que se está haciendo es un histograma de una variable binomial cuya media es 0.70." Esta respuesta confirma su capacidad para identificar la distribución que se obtendría al tomar muestras independientes de una distribución Bernoulli. Es importante destacar que la distribución resultante debe ser ponderada o normalizada para una comprensión completa. Destacamos que Raúl reconoció que el valor esperado es 0.7, lo que demuestra un razonamiento matemático en su respuesta.

En cuanto al ítem 9, se categorizó como "otro" debido a la imposibilidad de clasificarla según los niveles establecidos en el marco. En su respuesta, Raúl interpreta la pregunta como una prueba de hipótesis de cociente de verosimilitud generalizada, que pertenece al ámbito de las técnicas de inferencia bayesiana, aspecto que no está considerado en este estudio. Sin embargo, es importante señalar que el intervalo que Raúl elige para ubicar la muestra del problema, dentro del contexto de las pruebas informales de significación, no es la opción que permitiría rechazar la hipótesis de que en Europa existe una media superior en la esperanza de vida.

Tabla 5.2

Niveles de las respuestas por pregunta y alumno

	Población			Muestra		Muestreo repetido		Distribución muestral	
	1	2	3	4	5	6	7	8	9
Raúl	P	P	P	P	P	P	P	F	O
Susana	P	P	P	P	P	P	P	I	I
Javier	F	I	F	O	F	P	P	P	P
Daniel	P	P	P	P	I	P	F	P	P

Concepciones: P = Práctica; I = Informal, F = Formal, O = Otro tipo de respuestas

Respuestas por parte de la estudiante Susana

Susana proporcionó respuestas clasificadas en el nivel práctico en los ítems del 1 al 7 que se refieren a los conceptos de población, muestra y muestreo repetido, y como informales en los ítems 8 y 9 que se refieren a distribuciones muestrales. Ninguna de sus respuestas fue clasificada como formal. En el ítem 2, se le pidió a la estudiante elegir una muestra entre varios incisos, y ella eligió a y c, que son ejemplos concretos. Argumenta que elige estas opciones porque considera que "los elementos por sí mismos no revelan información alguna, simplemente existen". Desde su perspectiva, lo que relaciona a los elementos de dichos incisos como una población es su carácter empírico, lo que muestra una visión práctica de la población según el marco de desarrollo.

Tablas 5.3

Clasificación de las respuestas de Susana

	Población			Muestra		Muestreo Repetido		Distribución Muestral	
Ítems	1	2	3	4	5	6	7	8	9
Susana	P	P	P	P	P	P	P	I	I

En cuanto al concepto de muestra, en el ítem 5 se le pide elegir la opción que coincide con su concepto de población. La estudiante elige el inciso donde se realiza una extracción aleatoria de n elementos de un conjunto de m elementos y argumenta que una muestra no puede ser cualquier subconjunto, sino uno que lo represente bien, pero no explica exactamente qué significa que una muestra sea representativa ni cómo se consigue.

En el ítem 7, Susana interpreta que la opción a) se refiere a extraer una muestra amplia de 200 fichas y calcular su proporción, lo cual se alinea con la perspectiva aditiva del muestreo según nuestro marco.

Por otro lado, en el ítem 9, Susana selecciona una respuesta que no estaba listada entre las opciones, señala que los valores atípicos se ubican en los extremos de la distribución, admite no recordar el procedimiento exacto para realizar los cálculos que respalden su elección.

Respuestas por parte del estudiante Javier

Tablas 5.4

Clasificación de las respuestas de Javier

	Población			Muestra		Muestreo repetido		Distribución muestral	
	1	2	3	4	5	6	7	8	9
Javier	F	I	F	O	F	P	P	P	P

Javier se destaca por tener el mayor número de respuestas formales, con un total de tres, mientras que obtuvo cuatro respuestas clasificadas como nivel práctico, una respuesta informal y una respuesta que no encajaba en ninguna categoría específica ("otro").

En su respuesta al ítem 1, Javier definió una población como "un conjunto con su respectiva distribución de probabilidad". A pesar de la concisión, clasificamos su respuesta en el nivel matemático/abstracto, pues es una definición de naturaleza matemática. En el ítem 2, el estudiante selecciona las opciones b y c. La primera, opción b, es una definición matemática que asocia la población con una variable. La segunda, opción c, es un ejemplo específico, lo que sugiere que el estudiante contempla ambas representaciones. Él justifica su elección declarando:

Elegí las dos anteriores, porque a estas opciones a través de ciertas condiciones podemos proponer que tienen cierta distribución para cada elemento (tal vez variable) de la población y que cada una de estas son independientes e idénticamente distribuidas.

Esta argumentación evidencia que el estudiante relaciona el concepto de población con los conceptos de variable, independencia y distribución, lo que nos permite clasificar su nivel de entendimiento como transicional o abstracto. Sin embargo, conviene notar que no seleccionó la opción que directamente incluye la definición de una variable aleatoria general y su distribución; por esta razón, la ubicamos entre transicional y abstracto.

En el ítem 4, el estudiante Javier describe una muestra como "Una población con su respectiva distribución". Aunque esta concepción es intrigante, no está contemplada dentro de nuestro marco de desarrollo de razonamiento. La distinción entre "población" y "muestra",

y entre estos términos y "distribución", parece ser confusa en la respuesta del estudiante. Dado que la respuesta del estudiante al ítem 4 no es suficiente para permitir un análisis exhaustivo, y no se ajusta a las categorías previamente definidas en nuestro marco, la clasificamos bajo la categoría de 'otra respuesta'. Esta categoría surge para reconocer respuestas únicas que, aunque no se ajustan a las clasificaciones previstas, podrían brindar información sobre el razonamiento estadístico por parte de los estudiantes.

En el ítem 5, el estudiante únicamente afirma que una muestra es "un conjunto de elementos con su distribución", sin brindar argumentos adicionales. En este caso, no podemos determinar si el estudiante simplemente está replicando los argumentos presentes en las opciones de respuesta.

En el ítem 7, el estudiante selecciona la opción de muestreo sin reemplazo, explicando que esta opción "Evita la posibilidad de que se lleguen a repetir fichas en el conteo", y añade que "el promedio de las frecuencias es una forma de estimar la proporción de las fichas". Esta perspectiva, que considera el muestreo sin reemplazo como una opción más confiable, se alinea con una concepción práctica del concepto.

Finalmente, en el ítem 8, relacionado con la distribución muestral, el estudiante selecciona la representación que, en su opinión, refleja la información de la distribución dada. Por un lado, elige la distribución con media en 0.3 y, por otro lado, la distribución que tiene las 2 modas en punto 3 y punto 7. Esta elección refleja la visión práctica del estudiante de esperar que la distribución muestral se asemeje a la población, argumentando: "Elegí la opción porque esta gráfica cubre la información de cómo está dada la frecuencia dados los datos del enunciado".

Respuestas por parte del estudiante Daniel

Daniel se destaca por sus respuestas predominantemente prácticas, con un total de ocho respuestas clasificadas en esta categoría. Sin embargo, también proporcionó una respuesta formal en el ítem 7. Sólo en el ítem 5 se identificó una respuesta informal. En particular para los ítems sobre población (1-3) el estudiante afirma contundentemente que una población está conformada únicamente por personas, reflejando su visión práctica del concepto.

Para los ítems sobre muestra, Daniel en la pregunta abierta del ítem 4 responde que las muestras son subconjuntos de la población. Sin embargo, en la siguiente (ítem 5), considera que la aleatoriedad de la muestra es un factor importante mostrando una concepción informal o intermedia.

Con respecto al muestreo repetido menciona que eligió un muestreo con reemplazo en el ítem 7, ya que éste permite conservar el espacio muestral, es decir, la probabilidad de elegir un elemento se mantiene de ensayo a ensayo y éste es un rasgo de una visión más formal del muestreo. Sin embargo, en la pregunta previa sobre muestreo repetido ítem 6 menciona que para él sólo es: “Tomar un subconjunto de la población varias veces”. Reflejando una visión material y finita, que está descrita por el marco como una concepción práctica.

Tablas 5.5

Clasificación de las respuestas de Daniel

	Población			Muestra		Muestreo repetido		Distribución muestral	
	1	2	3	4	5	6	7	8	9
Daniel	P	P	P	P	I	P	F	P	P

Evaluación del Cuestionario por Expertos

Introducción de que esta sección presenta otro recurso de validación es esta consulta con los expertos

Se recurrió a la participación de tres expertos en educación estadística para actuar como evaluadores y proporcionar su opinión sobre los conceptos del marco y el instrumento de toma de datos, con el propósito de validar y enriquecer el instrumento de toma de datos para el marco propuesto.

Las respuestas a cada valoración de los ítems fueron clasificadas en tres categorías: "no propone modificaciones", "aportó algún comentario" y "desaprueba el ítem". El cuadro de las repuestas se presenta en la tabla 5.2, brindando una visión general de las opiniones de los expertos con respecto a los ítems del instrumento.

De los 11 ítems evaluados, los tres expertos estuvieron de acuerdo en que no era necesario realizar cambios en los ítems 1, 7, 9 y 10. Sin embargo, en los ítems 2, 3, 4, 5, 6 y 8, los evaluadores sugirieron algunas modificaciones. Solamente en el ítem 11, el evaluador 1 opinó que no era pertinente, mientras que los otros dos evaluadores expresaron reservas al respecto. Los tres evaluadores coincidieron en que el cuestionario podría ser aplicable con modificaciones menores. A continuación, se detallan aspectos importantes de las observaciones realizadas por los evaluadores al cuestionario. La versión que los evaluadores revisaron se puede consultar en el apéndice C En la siguiente exposición describiremos las repuestas de los expertos y en la discusión comentaremos las respuestas que desprenden de sus observaciones.

Evaluador 1

En relación a los 11 ítems evaluados, el Evaluador 1 dio su aprobación a 7 de ellos, específicamente los reactivos 1, 3, 4, 5, 7, 9 y 10. Además, en el caso del ítem 2, el evaluador señaló que la opción de respuesta en el inciso c, que presenta a la población como un conjunto de secuencias de ceros y unos, podría plantear cierta dificultad para los estudiantes al momento de responder. De manera similar, en el inciso c del ítem 6, se indicó que el uso de la expresión “distribución de Bernoulli” resultaba demasiado abstracta para el nivel de comprensión de los estudiantes.

En relación al ítem 8, donde se muestran diferentes modalidades para llevar a cabo un muestreo repetido, el evaluador expresó que varias de las respuestas podrían considerarse como adecuadas, lo cual plantea la cuestión de si se espera que los estudiantes elijan una única opción. Con respecto a esta duda, sí es posible elegir más de una opción, no obstante, es poco probable.

Por otro lado, en cuanto al ítem 11, el evaluador consideró que no resulta pertinente en el contexto del estudio, dado que se trata de un ítem relacionado con pruebas de hipótesis, y se sugirió su reemplazo por otro que aborde el concepto de distribución muestral. Más adelante se discutirá esta sugerencia.

Evaluador 2

El evaluador 2 otorgó su aprobación a 9 de los 11 ítems evaluados, específicamente los reactivos 1, 2, 3, 5, 6, 7, 8, 9, 10. En relación con el ítem 1, que plantea una pregunta abierta sobre el concepto de población, el evaluador destaca la importancia del ítem dado que usualmente no se aborda en la literatura.

En cuanto al ítem 4, que también es una pregunta abierta sobre el concepto de muestra, el evaluador señala que, en el cuestionario dirigido a los estudiantes, la especificación de que se trata del contexto de la estadística podría estar limitando el rango de respuestas, ya que los estudiantes podrían tener concepciones distintas basadas en sus propias experiencias.

Con relación al ítem 11, el evaluador considera que la redacción de la pregunta puede resultar demasiado compleja para los estudiantes, y sugiere cambiarla para abordar el concepto de inusualidad de muestras. Además, propone añadir una pregunta abierta sobre el concepto de distribución muestral.

Tabla 5.6

Cuadro de categorías de las respuestas de los evaluadores al marco

Conceptos del Marco de Desarrollo Del Razonamiento											
	Población			Muestra			Muestreo Repetido			Distribución Muestral	
Ítem #	1	2	3	4	5	6	7	8	9	10	11
Evaluador 1	✓	?	✓	✓	✓	?	✓	?	✓	✓	X
Evaluador 2	✓	✓	✓	?	✓	✓	✓	✓	✓	✓	?
Evaluador 3	✓	✓	?	✓	?	?	✓	✓	✓	✓	?
Valoración	Los tres evaluadores consideran que el cuestionario es aplicable con correcciones menores										
✓: No propone modificaciones; ?: Aportó algún comentario; X: Desaprueba el ítem.											

Evaluador 3

El evaluador 3 otorgó su aprobación a 7 de los 11 ítems evaluados, específicamente los reactivos 1, 2, 4, 7, 8, 9, 10. En relación con los ítems 3 y 6, también destaca que el uso de la expresión “distribución de Bernoulli” podría resultar poco clara para el nivel de comprensión de los estudiantes.

En el ítem 5, el evaluador propone mejorar la redacción, pero no especificó algún detalle extra. Con relación al ítem 11, plantea la preocupación de que no todos los estudiantes podrían comprender la formulación de una prueba de hipótesis.

Con relación a las observaciones planteadas por los evaluadores, es evidente que han aprobado la gran mayoría de los ítems del cuestionario. Resulta relevante resaltar que el evaluador 2 ha destacado la ausencia de estudios en la literatura que se refieran a las concepciones de los estudiantes sobre la noción de población. Los autores de este trabajo estamos de acuerdo con esta observación pues la notamos en la revisión de la bibliografía.

En relación con los ítems 2 y 5, tanto el Evaluador 1 como el Evaluador 3 han expresado su preocupación acerca de la claridad y comprensión de las opciones de respuesta, las cuales presentan a la población como un conjunto de valores numéricos. Como respuesta a esta inquietud, se tomó la decisión de ser más explícitos en la redacción y proporcionar una explicación clara sobre el significado de los valores numéricos presentados en las opciones (ver tabla 5.7).

Tabla 5.7

Modificación del ítem 2

Versión original del ítem 2 (ver apéndice C)	Reformulación del ítem
c) El conjunto de todas las secuencias o arreglos de 0 y 1 de tamaño 10. Ejemplo: (1,0,1,0,1,0,0,0,1,0,1).	c) El conjunto de todas las secuencias o arreglos de “0” y “1” de tamaño 10, donde 1 = “Encestar”, 0 = “No encestar”. Ejemplo de una secuencia de dicho conjunto: (1,0,1,0,1,0,0,0,1,0,1).

En cuanto a los ítems 3 y 6, se ha identificado que también los evaluadores 1 y 3, consideran que la inclusión de la frase “distribución Bernoulli” en los incisos podría resultar demasiado abstracta para los estudiantes. Con el objetivo de abordar esta preocupación, se ha optado por presentar la distribución Bernoulli como el resultado de un experimento

aleatorio con una probabilidad asociada (ver tabla 5.8), buscando así una formulación más accesible y comprensible para los alumnos.

Tabla 5.8

Modificación del ítem 3

Versión original del ítem 3 (ver apéndice C)	Reformulación del ítem
c) La distribución de Bernoulli con $B = "1"$, $N = "0"$, $P(1)=500/1300$ y $P(0)=800/1300$	c) Los valores de la variable "El resultado de sacar una única ficha de la urna que solo tiene fichas negras y blancas". Es decir, $B = \{0, 1\}$, donde la probabilidad de sacar una ficha blanca (1) es de $500/1300$, mientras que la probabilidad de sacar una ficha negra (0) es de $800/1300$

Con relación al ítem 4, el Evaluador 2 ha planteado la posibilidad de no hacer explícito que se trata de una muestra en el contexto de la estadística. No obstante, se considera que la inclusión del complemento "en estadística" en la pregunta es pertinente, ya que se busca que los estudiantes enfoquen su respuesta desde una perspectiva estadística y no desde la vida cotidiana, siendo esta diferenciación relevante para los objetivos de la investigación.

Respecto al ítem 8, en el cual se solicita a los estudiantes que elijan entre un muestreo repetido con o sin reemplazo. El evaluador 1 destaca que más de una opción podría ser válida y plantea la incertidumbre sobre si se espera que los estudiantes se decanten por una opción en particular. Nuestra expectativa es que los estudiantes elijan la opción que les parezca más informativa. Si bien reconocemos la validez de ambos tipos de muestreo, con o sin reemplazo, es importante considerar el objetivo de generar distribuciones muestrales simuladas, donde el muestreo con reemplazo permite realizar este tipo de distribuciones, además, el muestro con reemplazo ofrece un desarrollo matemático muchas más amplio que el muestreo sin reemplazo. En otras palabras, se debe tener en cuenta que se pretende abordar el enfoque matemático de la generación de la distribución muestral. No obstante, los estudiantes con un nivel de abstracción práctico se suelen declinar por el muestreo sin reemplazo.

En el ítem 11, los tres evaluadores coinciden en que la presentación de una prueba de hipótesis puede resultar demasiado complicada para los estudiantes. Específicamente, el Evaluador 2 propone un ítem en el que se aborde el problema de identificar una muestra

inusual a partir de una población. Reconocemos que ese ítem tiene que ver con pruebas de hipótesis, pero es precisamente en éstas que cobra sentido el concepto de distribución muestral por lo que decidimos mantenerlo. Además, por los resultados de la exploración con estudiantes podemos observar que, aunque los estudiantes no recuerden el procedimiento de una prueba de hipótesis es posible que ellos utilicen la distribución muestral para elegir una respuesta, lo cual brinda información para el análisis de la concepción de distribución muestral.

Aplicación del Cuestionario a dos Modelos de Lenguaje de Inteligencia Artificial

La aparición reciente de modelos de inteligencia artificial plantea el desafío de cómo utilizarla para potenciar el trabajo académico. Nosotros tratamos de aprovechar este recurso para apoyar la validación del instrumento de toma de datos, y encontramos resultados valiosos.

Como parte de la validación del instrumento de recolección de datos es imperativo examinar su capacidad para capturar en las respuestas recabadas los diversos niveles de razonamiento propuestos en nuestro marco de desarrollo. En este contexto, hemos recurrido al análisis de las respuestas de dos Modelos de Lenguaje Grandes (LLM por sus siglas en inglés): el GPT-3.5 y el GPT-4. Los modelos de lenguaje GPT son una instancia de la Inteligencia Artificial (IA) con la capacidad de procesar y generar lenguaje, ofreciendo una interacción similar a la que tendría lugar entre humanos. Se reconoce que estos modelos, producidos por la empresa OpenAI, exhiben diferencias en sus niveles de razonamiento, donde el modelo GPT-4 obtiene un mejor desempeño con respecto al modelo GPT-3.5 en pruebas de conocimiento para humanos (OpenAI, 2023).

Respuestas al cuestionario por parte de los modelos

Las respuestas al cuestionario se clasificaron en cuatro niveles: (1) Práctico, (2) Informal o Transicional, (3) Formal o Abstracto y (4) Integral. En la Tabla 5.5 se muestra el cuadro resultante de esta clasificación.

Se observa, como se esperaba, que el modelo GPT-4 supera al modelo GPT-3.5 en el nivel de razonamiento en 11 de 12 respuestas a los ítems del cuestionario de acuerdo a nuestra clasificación, solo en el ítem 4 presentan una respuesta equivalente. Específicamente, el

modelo GPT-4 obtiene 10 de 12 respuestas en los niveles Formal e Integral, mientras que el modelo GPT-3.5 registra 11 de 12 respuestas en los niveles Práctico e Informal. Lo cual nos permite observar una diferencia clara en el nivel de razonamiento.

Modelo GPT-3.5

El modelo GPT 3.5 responde en el nivel básico a los ítems 1, 2, 3, 4, y 6. En el nivel informal o transicional podemos clasificar los ítems 5, 7, 9, 10, 11 y 12. Únicamente en el ítem 8 el modelo GPT-3.5 logró responder en el nivel formal del razonamiento de acuerdo al marco de desarrollo.

En particular en los ítems 1, 2 y 3 que se refieren al concepto de población, para el modelo GPT-3.5 lo que determina a la población son los objetos o individuos, cuestión descrita por el marco de desarrollo como una visión práctica.

Tabla 5.9

Clasificación de las respuestas a los ítems por parte de los modelos de lenguaje GPT-3.5 y GPT-4

Conceptos del Marco de Desarrollo Del Razonamiento												
Ítems	Población			Muestra			Muestreo Repetido			Distribución Muestral		
	1	2	3	4	5	6	7	8	9	10	11	12
Modelo de lenguaje												
GPT-3.5	P	P	P	P	I	P	I	F	I	I	I	P
GPT-4	I	F	PIF	P	PIF	PIF	F	F	F	PIF	F	F

P: Razonamiento práctico; **I:** Razonamiento informal; **F:** Razonamiento formal;
PIF: Razonamiento Integral.

Para los ítems relacionados con el concepto de muestra el modelo GPT-3.5 sostiene la visión Práctica en los ítems 4 y 6, argumentando que las muestras son subconjuntos de la población que deben ser representativos, pero descartando las posibilidades más abstractas que pueden tener las muestras. En el ítem 5, donde se pregunta por una muestra de 10 tiros libres a la canasta, el modelo elige la opción del conjunto de números enteros correspondiendo a una visión simbólica de la muestra, pero descartando las demás opciones argumentando que no podrían ser representativas, pero sin argumentar por qué la opción elegida sí lo es.

Para los ítems 7, 8 y 9, relacionados con el muestreo repetido, el modelo GPT-3.5 sí considera el muestreo repetido como algo determinante, pero no logra relacionarlo con los aspectos más abstractos que se relacionan con este tipo de muestreo, como podría ser la distribución muestral o la inferencia.

En los ítems sobre distribución muestral el modelo GPT-3.5 relaciona la distribución con el proceso de inferencia, pero muestra tener confusiones para identificar si se está hablando de una distribución empírica, simulada o teórica. También, al momento de realizar cálculos u operaciones para responder suele tener errores en los procedimientos. Por ejemplo, en el ítem 12, donde se pregunta acerca de una prueba de significación, a pesar de tener la claridad del nivel de significación del 5%, no logra dar con un procedimiento para usar ese nivel de significación, junto con la distribución dada, para poder responder a la pregunta.

Modelo GPT-4

El modelo GPT-4 muestra diferentes niveles de respuesta en relación a los ítems del cuestionario. En el nivel práctico, se ubica solo al ítem 4, mientras que en el nivel informal solo se clasificó el ítem 1. En el nivel formal, están los ítems 2, 7, 8, 9, 11 y 12. En el nivel integral, se clasifican los ítems 3, 5, 6 y 10.

En relación al concepto de población, el modelo GPT-4 muestra diferentes representaciones. En el ítem 1, relaciona la población con objetos y medidas, lo cual es una visión informal. En el ítem 2, donde se menciona un experimento de lanzamiento de tiros libres, el modelo elige una opción más abstracta que corresponde al conjunto de valores con su respectiva distribución. En el ítem 3, donde se presentan diferentes opciones de población en el contexto de la extracción de fichas de una urna, el modelo elige la opción práctica, pero tiene en cuenta que las otras opciones también pueden ser útiles en contextos de simulación o estudios teóricos, lo cual muestra una visión amplia y un razonamiento integral.

En cuanto al concepto de muestra (ítems 4, 5 y 6), el modelo relaciona las muestras con subconjuntos de la población. En los ítems donde se selecciona una muestra, ya sea en el contexto de los tiros libres o de la extracción de fichas, el modelo explica que la elección depende del tipo de estudio y la elección correcta puede ser tanto práctica como abstracta, según las intenciones y el contexto, lo cual muestra una visión integral del concepto de muestra.

En los ítems relacionados al muestreo repetido, el modelo reconoce que el muestreo con reemplazo permite hacer estimaciones y relaciona el muestreo con la inferencia. Sin embargo, no menciona los aspectos matemáticos del muestreo sin reemplazo y sus implicaciones en el cálculo de probabilidades. Es crucial señalar que, en el muestreo sin reemplazo, la probabilidad de seleccionar cada elemento varía tras cada extracción, lo cual afecta las probabilidades de selección subsiguientes y, por ende, las inferencias estadísticas derivadas de dichas muestras. Este detalle matemático es fundamental para comprender las diferencias entre los dos métodos de muestreo y sus implicaciones en estudios estadísticos. En el ítem 9, donde se realiza una estimación proporción de fichas a partir de un muestreo repetido de una urna, el modelo elige el muestreo con reemplazo y destaca que la media de las muestras es un mejor estimador que la moda, lo cual muestra un razonamiento abstracto y diferenciado del modelo GPT-3.5.

En relación a los ítems sobre la distribución muestral, el modelo GPT-4 exhibe claramente la relación entre la distribución muestral y el teorema central del límite. Además, tiene la capacidad de realizar una prueba de significación informal utilizando la distribución muestral simulada presentada en el ítem como referencia. A diferencia del modelo GPT-3.5, que comprende el problema, pero tiene dificultades para realizar los cálculos y dar una respuesta adecuada. La capacidad del modelo GPT-4 de realizar pruebas de significación informales y relacionar la distribución muestral con el patrón de variación y el teorema del límite central, muestra un nivel de razonamiento integral con respecto a la noción de distribución muestral.

Segunda exploración con el instrumento con profesores de CCH

En un taller intersemestral de un bachillerato público en la ciudad de México con 11 profesores cuyo tema fue “sobre la enseñanza y aprendizaje de pruebas de significación con ayuda de tecnología”. Se les pidió a los profesores que respondieran de manera individual el cuestionario correspondiente al apéndice E. Presentamos de forma general los resultados por cada grupo de ítems relacionado a cada uno de los conceptos del marco, para una visión panorámica ver tabla 5.10.

Once de los participantes respondieron el cuestionario. Las respuestas se clasificaron de acuerdo con el marco propuesto y la distribución de las frecuencias por niveles y preguntas

se muestra en la tabla 5.6. Como se puede observar (resaltado por sombreado en tonos de gris) hubo variabilidad en el nivel de las respuestas tanto por profesor como por pregunta.

En los ítems sobre población y muestra, predominan las respuestas de tipo "Práctico", indicando que los participantes suelen basarse en situaciones concretas. Sin embargo, en los ítems sobre muestreo repetido, hay un equilibrio entre razonamientos prácticos y formales. Por último, en los ítems sobre distribución muestral, el razonamiento "Formal" es el más predominante, reflejando un enfoque más matemático por parte de los participantes

Tabla 5.10

Niveles de respuesta de cada profesor por ítem

Profesor	Población				Muestra		Muestreo Repetido			Distribución Muestral		
	1	2	3	4	5	6	7	8	9	10	11	12
1	I	F	F	P	F	F	F	F	F	F	F	I
2	I	F	I	P	I	I	F	F	F	I	F	F
3	I	F	P	P	P	P	F	P	F	I	I	P
4	P	P	P	P	P	P	I	I	P	F	F	I
5	I	I	F	P	I	F	F	I	F	I	F	P
6	P	I	F	P	I	I	P	I	P	P	F	I
7	P	P	P	P	P	P	I	P	F	I	F	I
8	P	P	P	P	P	P	P	F	F	F	P	F
9	P	F	F	P	F	F	I	P	F	F	F	P
10	F	F	F	P	P	P	P	I	I	P	F	P
11	P	P	P	P	P	P	F	F	F	F	F	I

Triangulación de las respuestas

Aplicamos un tratamiento a la matriz que representa la tabla 5.6 en el que reordenamos las filas (profesores) y las columnas (ítems) del mayor al menor número de respuestas en el nivel práctico; es decir, triangulamos la matriz; con base en ella se pudieron identificar grupos de ítems (A, B, y C) y grupos de profesores (X, Y, Z) como se muestra en la tabla 5.11.

Tabla 5.11

Triangulación y agrupación de los niveles de respuestas por profesor vs ítem

		A		B			A	C					
		4	1	5	6	3	2	12	7	8	10	9	11
X	8	P	P	P	P	P	P	F	P	F	F	F	P
	4	P	P	P	P	P	P	I	I	I	F	P	F
	7	P	P	P	P	P	P	I	I	P	I	F	F
	11	P	P	P	P	P	P	I	F	F	F	F	F
Y	10	P	F	P	P	F	F	P	P	I	P	I	F
	3	P	I	P	P	P	F	P	F	P	I	F	I
	6	P	P	I	I	F	I	I	P	I	P	P	F
	9	P	P	F	F	F	F	P	I	P	F	F	F
Z	5	P	I	I	F	F	I	P	F	I	I	F	F
	2	P	I	I	I	I	F	F	F	F	I	F	F
	1	P	I	F	F	F	F	I	F	F	F	F	F
P		11	6	6	6	5	4	4	3	3	2	2	1

Los profesores del grupo X muestran una tendencia a responder en el nivel práctico a todas las preguntas del grupo B. Sin embargo, no responden de la misma manera a las preguntas del grupo C, optando por respuestas formales o informales. Esto indica que tener respuestas prácticas en temas de población y muestra no garantiza respuestas similares en temas de muestreo repetido y distribución muestral. Por otro lado, las respuestas de los profesores del grupo Y en el grupo de ítems B están relacionadas con sus respuestas en el grupo C, reafirmando la falta de correlación entre los niveles de respuesta de ambos grupos de preguntas. Por último, los profesores del grupo Z consistentemente evitan respuestas prácticas en los grupos B y C, eligiendo en su lugar respuestas de niveles informal o formal.

Discusión

Vygotsky (1995) identifica el conocimiento espontáneo como aquel adquirido naturalmente a través de la vida diaria y las interacciones culturales. Este conocimiento espontáneo se basa en la experiencia cotidiana y en la comunicación común. Además, sirve de cimiento para el desarrollo del conocimiento científico formal y sistemático. Por otro lado, el conocimiento científico se caracteriza por ser adquirido mediante métodos educativos formales, la

investigación rigurosa y la reflexión crítica, permitiendo una comprensión y transformación más detallada y exacta del mundo.

La visión de Vygotsky puede servir un marco para entender cómo los estudiantes y profesores se acercan a conceptos estadísticos. Mientras que los conceptos de población y muestra se adquieren de manera espontánea en la convivencia social, sin necesidad de relacionarlos con las matemáticas. En cambio, el muestreo repetido y la distribución muestral son conceptos que sólo se aprenden en el curso de estadística y, en estos, se aborda de manera formal.

En los cursos se enfatiza menos en los rasgos formales de la definición de población y muestra que en las definiciones de muestreo repetido y distribución muestral. Esto podría explicar que se elijan las opciones prácticas en las preguntas del grupo B y se elijan niveles formales en algunas preguntas del grupo C. Estos resultados sugieren, al menos en el caso del grupo de profesores del grupo X, que los niveles de respuesta están influidos por el tipo de concepto (espontáneo o científico) contenido en el problema. En el otro extremo, por lo que concierne al grupo Z de profesores, parece que responder los ítems del grupo B de manera informal o formal implica responder en los mismos niveles las preguntas del grupo C, pues este grupo no responde ninguna pregunta al nivel práctico.

Capítulo VI. Conclusiones

En esta investigación se han llevado a cabo cinco formas de producir datos para validar el instrumento que, a su vez, permiten sostener que el marco de desarrollo propuesto refleja una realidad acerca de las concepciones de los estudiantes sobre los conceptos de muestreo abordados en este estudio. En cada una de estas formas de validación han surgido observaciones que permiten modificar el instrumento, algunas de ellas las hemos implementado y otras quedan para su futuro mejoramiento. Identificamos cada una de las formas con los siguientes identificadores:

- Entrevista con la profesora
- Exploración con estudiantes
- Consulta con expertos
- Uso del ChatGPT 3.5 y 4.0
- Exploración con profesores

En el capítulo previo se han expuesto con detalle los resultados principales de cada uno de estos acercamientos al problema de la validez del instrumento y el marco. Consideramos que cada uno de ellos aporta elementos para afirmar que el marco de desarrollo refleja efectivamente procesos de evolución de las concepciones por los que pasan los estudiantes durante su aprendizaje del tema, en particular, de los cuatro conceptos que hemos destacado. Enseguida, resumiremos lo que consideramos la aportación de cada uno de tales acercamientos. El orden que tienen en la lista anterior es con el que se llevaron a cabo, sin embargo, en la exposición que sigue alteramos en un orden que va de las más relevantes (en función de las evidencias) a las exploraciones menos relevantes.

La exploración con profesores

Proporcionó una mayor cantidad de datos, destacando la utilidad del instrumento para discernir distintos niveles de abstracción en las concepciones de los docentes. A través del análisis, se identificaron subgrupos de profesores con patrones de respuesta variados, lo que permitió formular hipótesis sobre los niveles de desarrollo de sus concepciones. Además, se constató que los conceptos de población y muestra son de naturaleza diferente a los conceptos de muestreo repetido y distribución muestral, debido a la forma en que son aprendidos.

El uso de GPT-3.5 y GPT-4

Esta fase proporcionó otro tipo de evidencia sobre el instrumento. Los resultados muestran que mientras las respuestas de GPT-3.5 se clasificaron en el nivel práctico, las de GPT-4 se identificaron como matemáticas, lo cual confirma que el cuestionario puede diferenciar adecuadamente los niveles de desarrollo conceptual.

Además, las respuestas de GPT-4 resaltan la importancia de considerar un nivel integral en nuestro análisis. Este nivel refleja cómo los tres niveles de razonamiento — práctico, informal y matemático— pueden ser aplicables dependiendo del contexto del problema y las circunstancias específicas en las que se encuentra el resolutor. La inclusión de este nivel integral permitiría un enfoque más amplio que reconoce que distintas perspectivas pueden ser válidas y útiles para abordar y resolver problemas estadísticos, adaptándose a las múltiples situaciones en las que se presenta o utiliza el concepto.

La consulta con expertos

Resulta relevante porque dieron su visto bueno al cuestionario como un instrumento para evaluar el marco. No obstante, sus evaluaciones fueron poco informativas al respecto, pero detectaron detalles menores para mejorar o incluso cambiar algunos, muy pocos, ítems del cuestionario.

La exploración con estudiantes

Esta fase arrojó datos limitados debido a que solo cuatro completaron el cuestionario. A pesar de esto, sus respuestas muestran que aún después de un primer curso universitario de estadística, tienden a responder en un nivel práctico. Este comportamiento muestra que los estudiantes comprenden las preguntas y aplican los conocimientos adquiridos en su curso. Sin embargo, para validar la capacidad del instrumento de detectar diferentes niveles de desarrollo conceptual, es necesario aplicarlo a una muestra más amplia de estudiantes con características similares a las de los participantes actuales.

La entrevista con la maestra

Su opinión sobre la importancia de desarrollar un marco de desarrollo, aunque no se obtuvo información directa sobre la validez del instrumento utilizado. Ella consideró que el marco es útil para orientar la enseñanza de los conceptos clave de muestreo, evidenciando pertinencia para la enseñanza. Sin embargo, también sugirió que el marco debería ampliarse para incluir conceptos adicionales como aleatoriedad, técnicas de muestreo y probabilidad.

Limitaciones

Las principales limitaciones del estudio fueron el tamaño reducido de la muestra de estudiantes debido a la pandemia, lo que limitó la riqueza de los datos obtenidos. El instrumento empleado para evaluar a los expertos resultó demasiado restrictivo y pudo haber sido reemplazado por una entrevista para obtener respuestas más profundas. La falta de familiaridad de la maestra entrevistada con los métodos de investigación en educación estadística pudo influir en su percepción sobre el marco de desarrollo evaluado. En la exploración con profesores, se identificó la necesidad de contar con información sobre sus antecedentes académicos para formular hipótesis más precisas sobre sus niveles de respuesta.

Acciones para investigaciones futuras:

1. Se propone una reelaboración del instrumento para evaluar distintos contextos en los que se pueden aplicar los conceptos estadísticos. Esto incluiría categorizar las preguntas bajo tres contextos específicos:

- Social, tecnológico o científico
- Manipulativos: Juegos de azar, ruletas y urnas (modelos pseudoconcretos).
- Matemático.

Esta clasificación permitiría obtener más información de cómo los estudiantes razonan en diversos escenarios prácticos, intermedios y teóricos.

2. Aumentar el tamaño de la muestra de estudiantes.

3. Hacer entrevistas con los expertos para profundizar en la visión de los aspectos observados por el instrumento y obtener conocimientos más matizados sobre su aplicabilidad y relevancia.

Referencias

- Ainley, J., Gould, R. y Pratt, D. (2015). Learning to reason from samples: commentary from the perspectives of task design and the emergence of 'Big data'. *Educational Studies in Mathematics* 88(3), 405–412
- Bakker, A. (2018). *Design Research in Education. A practical Guide for Early Career Researchers*, Routledge.
- Batanero, C., López-Martín, M., Gea, M., & Arteaga, P. (2018). Conocimiento del contraste de hipótesis por futuros profesores de Educación Secundaria y Bachillerato. *PUBLICACIONES*, 48(2), 73–95. <https://doi.org/10.30827/publicaciones.v48i2.8334>
- Ben-Zvi, D. (2004). Reasoning about variability in comparing distributions. *Statistics Education Research Journal*, 3(2), 42–63.
- Biggs, J. & Collis, K. (1982). *Evaluating the Quality of Learning: the SOLO Taxonomy*. New York: Academic Press.
- Biggs, J. & Collis, K. (1991). Multimodal learning and the quality of intelligent behavior. In H. Rowe (Ed.), *Intelligence, Reconceptualization and Measurement* (pp. 57-76). New Jersey: Laurence Erlbaum Assoc.
- Case, C., & Jacobbe, T. (2018). A framework to characterize student difficulties in learning inference from a simulation-based approach¹. *Statistics Education Research Journal*, 17(2). <https://doi.org/10.52041/serj.v17i2.156>
- Chance, B., del Mas, R., y Garfield, J. (2004). Reasoning about Sampling Distributions. In D. Ben-Zvi y Garfield (Eds.). *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 295–323). Springer.

- Cobb, G. W. (2007). The Introductory Statistics Course: A Ptolemaic Curriculum? *Technology Innovations in Statistics Education*, 1(1). <http://dx.doi.org/10.5070/T511000028> Retrieved from <https://escholarship.org/uc/item/6hb3k0nz>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge. <https://doi.org/10.4324/9780203029053>
- Confrey, J. (1990). A Review of the research on student conceptions in mathematics, science, and programming. *Review of Research in Education*, 16, 3-56.
- Dreyfus, T. (1991). The nature of advanced mathematical thinking. En D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 25-41). Kluwer.
- Fujii, T. (2020). Misconceptions and Alternative Conceptions in Mathematics Education. En S. Lerman (ed.), *Encyclopedia of Mathematics Education* (pp. 625-627).
- Freudenthal, H. (1982). Fiabilidad, validez y pertinencia. Criterios de la investigación sobre el aprendizaje de la matemática, *Educational Studies in Mathematics* 13(4) 395--408.
- García, V. (2017) *Diseño de una Trayectoria Hipotética de Aprendizaje para la Introducción y Desarrollo del Razonamiento sobre el Contraste de Hipótesis en el Nivel Medio Superior*. Tesis de doctorado no publicada. Departamento de Matemática Educativa, Cinvestav-IPN.
- Hogg, R. V., Carig, A. T. (1970). *Introduction to Mathematical Statistics* (3rd edition). Macmillan.
- Jones, G. A., Langrall, C. W., Thornton C. A. y Mogill, A. T. (1997). A framework for assessing and nurturing young children's thinking in probability. *Educational Studies in Mathematics* 32(2), 101-125
- Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B y Putt, I. J. (2000). A framework for characterizing children's statistical thinking, *Mathematical Thinking and Learning*, 2(4), 269-307

- Langrall, C.H., Makar, K., Nilsson, P., Shaughnessy, J. M. (2017). Teaching and Learning probability and statistics: An integrated perspective. En J. Cai (Ed.), *Compendium for Research in Mathematics Education*, National Council of Teachers of Mathematics.
- Lobato, J., & Walters, C. D. (2017). A Taxonomy of Approaches to Learning Trajectories and Progressions. En J. Cai (Ed.), *The Compendium for Research in Mathematics Education: National Council of Teachers of Mathematics*.
- National Research Council. (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8*. Washington, DC: The National Academies Press. [doi: 10.17226/11625](https://doi.org/10.17226/11625).
- Noll, J., y Shaughnessy, M. (2012). Aspects of students' reasoning about variation in empirical sampling distributions. *Journal for Research in Mathematics Education*, 43(5), 509–556.
- OpenAI. (2023). GPT-4 Technical Report. arXiv: 2303.08774.
- Reading, C., y Reid, J. (2006). An emerging hierarchy of reasoning about distribution: From a variation perspective. *Statistics Education Research Journal*, 5(2), 46–68.
- Reid, J., y Reading, C. (2008). Measuring the development of students' consideration of variation. *Statistics Education Research Journal*, 7(1), 40–59
- Saldanha, L. y Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics* 51(3), 257-270.
- Saldanha, L. y Thompson, P. (2007). Exploring connections between sampling distributions and statistical inference: an analysis of students' engagement and thinking in the context of instruction involving repeated sampling. *International Electronic Journal of Mathematics Education*, 2(3).
- Schwartz, D. L., Goldman, S. R., Vye, N. J., Barron, B. J. (2000). Aligning everyday and mathematical reasoning: The case of sampling assumptions. En S. P. Lajoie (Ed.), *Reflections on Statistics*.

- Learning, Teaching, and Assessment in Grades K-12* (pp- 233-273) Lawrence Erlbaum Associates Publishers.
- Silvestre, E. (2019). *Actividades sobre distribuciones muestrales utilizando simulaciones en el bachillerato: un experimento de diseño*. Tesis de doctorado no publicada. Departamento de Matemática Educativa, Cinvestav-IPN.
- Silvestre C., Sánchez, E., e Inzunza, S. (2022). El razonamiento de estudiantes de bachillerato sobre el muestreo repetido y la distribución muestral empírica. *Educación Matemática*, 34(1), 100–130. <https://doi.org/10.24844/em3401.04>
- Sóbol, I. M. (1976). *El Método de Montecarlo*. Editorial MIR.
- Spangler, D.A., & Williams, S.R. (2019). The Role of Theoretical Frameworks in Mathematics Education Research. In K.R. Leatham (Ed.), *Designing, Conducting, and Publishing Quality Research in Mathematics Education* (pp. 3-18). Springer, Cham. https://doi.org/10.1007/978-3-030-23505-5_1
- Stanley, D. E., & Nyrupe, R. (2020). Strategies in abduction: Generating and selecting diagnostic hypotheses. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, 45 (2), 159–178. <https://doi.org/10.1093/jmp/jhz041>
- van Dijke-Droogers, M., Drijvers, P. y Bakker, A. (2019). Repeated sampling with a black box to make informal statistical inference accessible. *Mathematical Thinking and Learning*, 22 (2), 116–138.
- Watson, J. M. y Moritz, J. B. (2000a). Development of understanding of sampling for statistical literacy. *Journal of Mathematical Behavior* 19 (1) 109–136.
- Watson J. M. y Moritz, J. B. (2000b). Developing Concepts of Sampling. *Journal for Research in Mathematics Education*, 31(1), 44-70.

Watson, J. M., Callingham, R. A., y Kelly, B. A. (2007). Students' appreciation of expectation and variation as a foundation for statistical understanding. *Mathematical Thinking and Learning*, 9(2), 83–130.

Apéndice A: Instrumento presentado en la entrevista con la profesora experta en estadística.

1. ¿Qué ejemplo coincide mejor con el concepto que tienes de población estadística?

- f) Los estudiantes de la facultad.
- g) Todos los posibles resultados de lanzar una moneda 10 veces.
- h) Una caja con 500 pelotas negras y 800 blancas.
- i) Una distribución generada por una variable aleatoria.
- j) Otra. Explica tu respuesta

2.- Se busca ver cómo se distribuyen 500 muestras de tamaño n. En cada muestra se mide la frecuencia de elección del refresco coca cola sobre el refresco pepsi. Solo hay esas dos opciones. ¿Cuál es la población del problema?

- a) Las personas que toman coca cola
- b) Distribución Bernoulli con dos resultados: éxito = prefiere coca cola, fracaso = no prefiere coca, y con función de probabilidad:

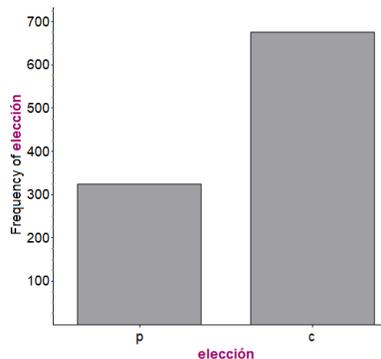
$$f(x)=\begin{cases} p, & \text{si } x = \text{éxito} \\ 1 - p, & \text{si } x = \text{fracaso} \end{cases}$$

- c) Las personas que toman refresco.
- d) El conjunto de todas las muestras.
- e) Otra. Explica tu respuesta.

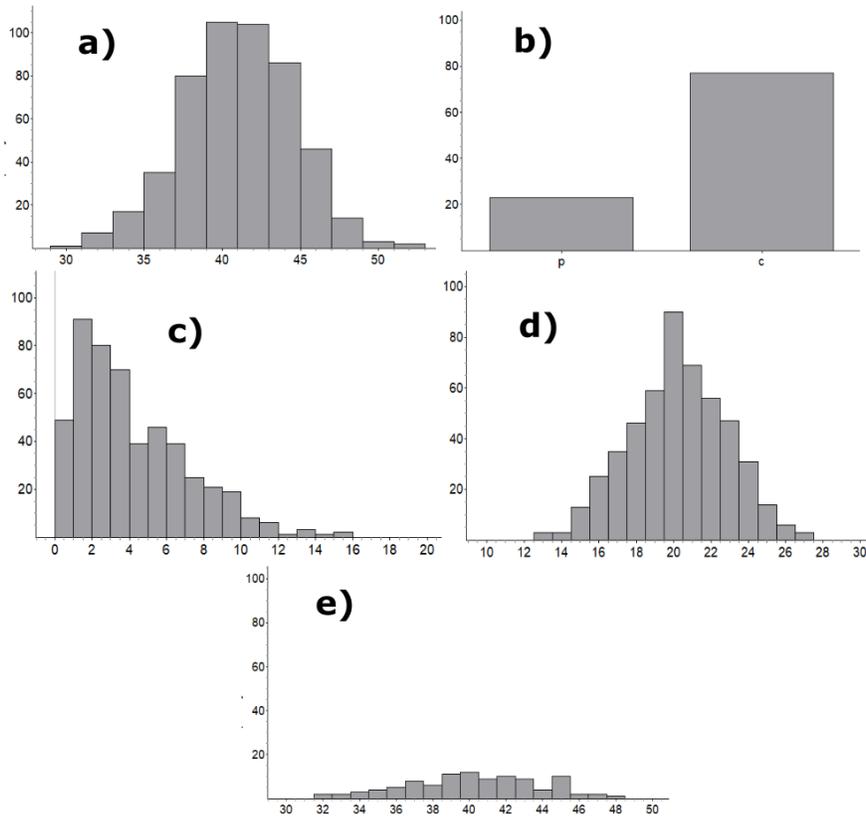
3.- ¿Qué ejemplo coincide mejor con el concepto que tienes de muestra estadística?

- a) Todos los casos de un lanzamiento de dos dados.
- b) Un conjunto de valores de una variable aleatoria elegidos de acuerdo con su distribución.
- c) Los alumnos de un grupo en particular de una preparatoria.
- d) Cualquier subconjunto de individuos de la población.
- e) Otra. Explica tu respuesta.

4.- Observa la siguiente imagen. Corresponde a la distribución en la preferencia en elección de coca cola o pepsi de 1000 personas, 325 personas eligieron pepsi mientras que 675 eligieron coca cola.



Se realizan 500 muestras de tamaño 30 de la distribución anterior, en cada muestra se registra frecuencia de la coca cola. ¿Cuál de las siguientes gráficas representa la distribución de frecuencia de las 500 muestras?



4.- Se tiene una caja con fichas negras y blancas, pero se desconoce la cantidad exacta de cada color, solo se sabe que suman 2000 fichas.

¿Cuál crees que sería la mejor forma de estimar la proporción de fichas negras si se te permite sacar 20 muestras de 10 fichas?

- a) Sacar una muestra, anotar el número de fichas negras, dejar a un lado la muestra y volver a sacar una muestra nueva. Al final, sumar el resultado de todas las muestras obtenidas y dividir entre 200.
- b) Registrar el número de fichas negras por muestra. Se quitan del proceso las muestras obtenidas. Se calcula la proporción de fichas negras en cada muestra y al final se hace un promedio con todas las proporciones.
- c) Tomar una muestra anotar la cantidad de fichas negras, se regresan las muestras a la caja con las demás fichas. Así hasta obtener las 20 muestras. Elegir como una buena estimación de la proporción la muestra que más veces apareció con cierta proporción de fichas negras.

- d) Tomar una muestra anotar la cantidad de fichas negras, se regresan las muestras a la caja con las demás fichas. Al final, se suma el resultado de todas las muestras obtenidas y se divide entre 200.
- e) Otra. Explica tu respuesta.

5. Existe la hipótesis de que en una elección estatal entre dos candidatos A y B, existe una mayor preferencia por el candidato B. Para ello se realiza una simulación por computadora donde se simulan 500 muestras de tamaño 1000 de proporción de votaciones para los candidatos A y B. La simulación se realiza suponiendo que la probabilidad de elección para cualquiera de los dos candidatos es la misma igual, es decir 0.5 o 50%.

Posteriormente, se realiza una encuesta de forma aleatoria a 1000 personas y se encuentra una proporción de votos para el candidato B de 0.78 o 78%. Con base en la simulación, la probabilidad de obtener una proporción así o mayor sería de .037 o 3.7%.

Lo cual conduce a aceptar la hipótesis de que la hay una mayor preferencia por el candidato B. Ya que la probabilidad de una muestra así es baja suponiendo que ambos tienen la misma probabilidad de ser elegidos.

Con base en el experimento anterior, cuál opción crees que refleja mejor tu pensamiento.

- a) Se debe rechazar la hipótesis porque la simulación no está hecha con personas reales.
- b) Es razonable aceptar la hipótesis siempre y cuando se considere que existe la probabilidad de estar equivocado.
- c) Se debe rechazar la hipótesis porque la simulación no debe suponer que existe la misma probabilidad en la elección de un candidato.
- d) Es razonable aceptar la hipótesis, pero la encuesta requiere de más personas.
- e) Otra. Explica tu respuesta.

**Apéndice B: cuestionario usado en la exploración con alumnos
(segunda fase)**

1.- ¿Qué entiendes por una población en el contexto de la Estadística?

(Respuesta abierta)

2.- ¿Qué ejemplo puedes relacionar con el concepto que tienes de población? (Puedes elegir más de una opción).

- a) Los estudiantes de la facultad.
- b) Todos los posibles resultados de lanzar una moneda 10 veces.
- c) Una caja con 500 pelotas negras y 800 blancas.
- d) Una variable aleatoria con su respectiva distribución.
- e) Otra. Explica tu respuesta abajo.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior.

(Respuesta abierta)

3. Se quiere saber si la preferencia de elección entre los refrescos Pepsi y Coca-Cola es la misma o si existe mayor preferencia por uno de los dos. Con esta información: ¿Cuál podría ser la población de la cual se quiere obtener información? (Puedes elegir más de una opción).

- a) Las personas que toman coca cola
- b) Distribución Bernoulli con dos resultados: Éxito = prefiere coca cola, que ocurre con probabilidad p . Fracaso = no prefiere coca, que ocurre con probabilidad $1-p$.
- c) Las personas que toman refresco de cola
- d) Las personas que toman refresco
- e) Otra. Explica tu respuesta abajo.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior

(respuesta abierta)

4. ¿Qué entiendes por una muestra en el contexto de la estadística?

(respuesta abierta)

5. ¿Qué ejemplo coincide mejor con el concepto que tienes de muestra estadística? (Puedes elegir más de una opción)

- f) Todos los casos de un lanzamiento de dos dados.
- g) Un conjunto de n valores de una variable aleatoria elegidos al azar de acuerdo con su distribución probabilidad.
- h) Un grupo de n alumnos elegidos al azar de una preparatoria con M alumnos.
- i) Cualquier subconjunto de individuos de la población.

Otra. Explica tu respuesta abajo.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior.

(Respuesta abierta)

6. ¿Qué entiendes por un muestreo repetido?

(Respuesta abierta)

El siguiente problema es una versión modificada al que aparece en la quinta sección en (Chance et al. 2004)

7. Se tiene una caja con fichas negras y blancas, pero se desconoce la cantidad exacta de cada color, solo se sabe que suman 2000 fichas.

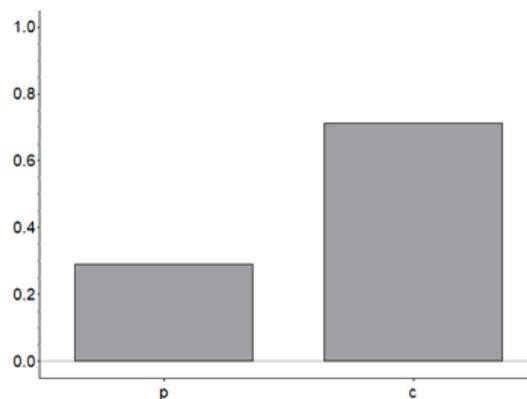
¿Cuál crees que sería la mejor forma de estimar la proporción de fichas negras si se te permite sacar 20 muestras de 10 fichas?

- a) Sacar una muestra, registrar la proporción de fichas negras, dejar fuera esa muestra de la caja y volver a sacar una muestra nueva, así sucesivamente. Al final, sumar la proporción de todas las muestras obtenidas y dividir entre 20.
- b) Sacar una muestra, registrar la proporción de fichas negras, dejar fuera esa muestra de la caja y volver a sacar una muestra nueva, así sucesivamente. Al final observar cual proporción apareció más veces.
- c) Sacar una muestra, registrar la proporción de fichas negras, regresar la muestra a la caja y volver a sacar una muestra nueva, así sucesivamente. Al final, sumar la proporción de todas las muestras obtenidas y dividir entre 20.
- d) Sacar una muestra, registrar la proporción de fichas negras, regresar la muestra a la caja y volver a sacar una muestra nueva, así sucesivamente. Al final observar cual proporción apareció más veces.
- e) Otra. Explica tu respuesta abajo.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior

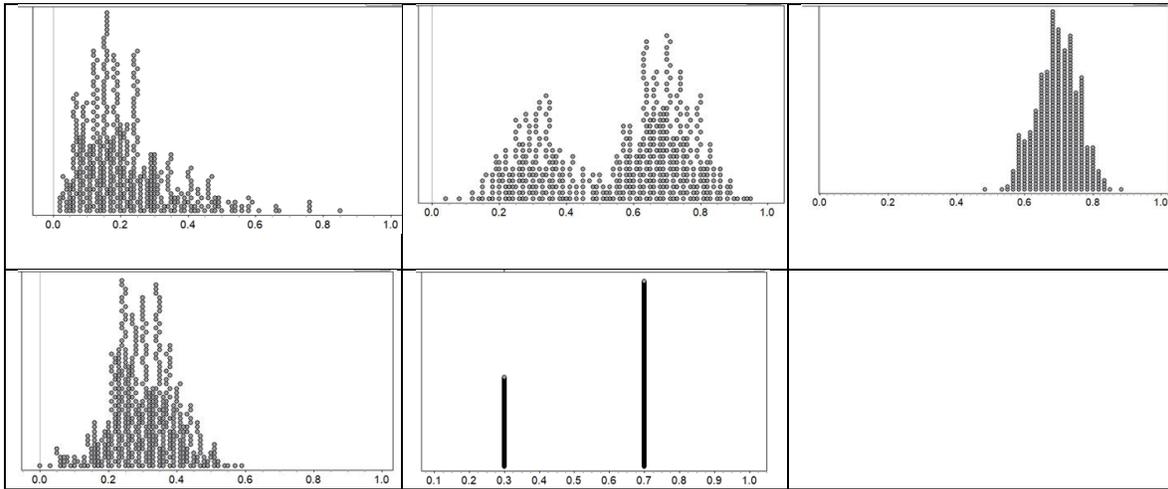
(Respuesta abierta)

8. La siguiente imagen corresponde a la distribución de la probabilidad de elegir entre el refresco Coca cola o Pepsi. Se elige Coca cola con una probabilidad de 0.7 y la probabilidad de Pepsi es de 0.3.



Se realizan 500 muestras independientes entre sí de tamaño 30 de la distribución anterior, en cada muestra se registra proporción de elección de coca cola. Es decir, se cuentan el número de elecciones de Coca-Cola y se divide entre 30, por lo que cada muestra arroja un número entre 0 y 1. }

Cada muestra es representada con punto de la gráfica y si tienen la misma proporción se apilan los puntos. ¿Cuál de las siguientes gráficas representa la distribución de frecuencia de la proporción de las 500 muestras?

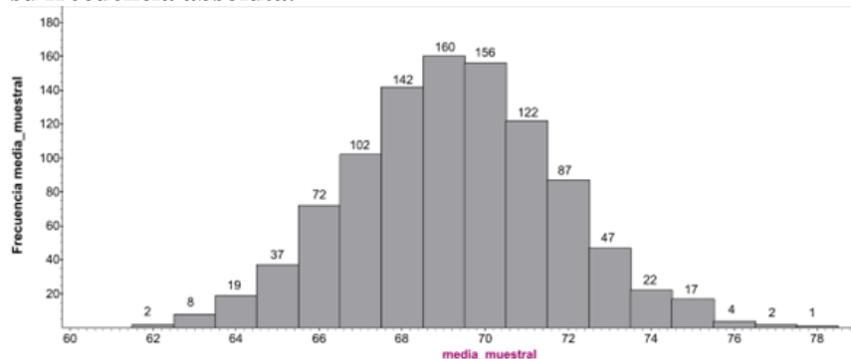


¿Por qué elegiste la(s) opción(es) en la pregunta anterior?

(Respuesta abierta)

El siguiente problema es una versión modificada al que aparece en (Batanero et al., 2018)

9. Se sabe que la esperanza de vida de los países se distribuye normalmente con una media 69.2 años y desviación estándar de 15. A partir de los datos mundiales se simula una muestra aleatoria de tamaño 35 y se obtiene su media. Este procedimiento se repite 1000 veces y se obtiene la distribución de frecuencias de las medias de las 1000 muestras obtenidas en la simulación. La siguiente gráfica muestra la distribución correspondiente. Cada número sobre las barras es su frecuencia absoluta.



Si se obtuviera una muestra real de países de Europa de tamaño 35 y se calculara su media, ¿aproximadamente a partir de qué valor del estadístico (en este caso la media) de la muestra real es significativo con un nivel de significación del 5%? Es decir, ¿a partir de qué valor del estadístico se podría rechazar la hipótesis de que los países de Europa tienen una esperanza de vida igual a la media mundial? Usa los valores de la distribución de arriba para responder la pregunta.

- Si el estadístico tiene un valor mayor a 69.2
- A partir de un valor 74, porque 74 o más tienen aproximadamente frecuencia del 5% del total.
- Si es igual 73, porque tiene frecuencia cercana al 5% del total.

- d) Si el estadístico cae en el intervalo de [63, 73], porque los fuera de ese intervalo extremos tienen frecuencias aproximadas a 5% del total.
- e) Otro valor. explica tu respuesta.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior
(Respuesta abierta)

Referencias:

Chance, B., del Mas, R., & Garfield, J. (2004). Reasoning about Sampling Distributions. In *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 295–323). Springer Netherlands.

Batanero Bernabeu, C., López-Martín, M. del M., Gea Serrano, M. M., & Arteaga Cezón, P. (2018). Conocimiento del contraste de hipótesis por futuros profesores de Educación Secundaria y Bachillerato. *PUBLICACIONES*, 48(2), 73–95.
<https://doi.org/10.30827/publicaciones.v48i2.8334>

Apéndice C: cuestionario sobre concepciones acerca de conceptos de muestreo

Apreciable profesional, le agradeceremos identificar los ítems y contestar marcando con una “X” en la tabla correspondiente, sobre la formulación, pertinencia y relevancia de cada ítem. Además, apreciaremos mucho sus observaciones, sugerencias o comentarios en los espacios previstos para ello.

Población objetivo. Se aplicará a estudiantes de bachillerato y universitarios (ingeniería, psicología, actuaría) que hayan llevado un primer curso de estadística.

Criterios:

Formulación: Juzgue el lenguaje y su claridad, utilizado en la presentación de cada ítem. Marque lo correspondiente en la columna de formulación, según los criterios “Adecuada”, “A mejorar” o “No adecuada”.

Pertinencia: Juzgue si el ítem es apropiado o congruente con lo que se espera evaluar, es decir, si las respuestas pueden proporcionar información sobre las categorías del marco. Refiérase en términos de “Pertinente”, “Con dudas” o “No pertinente”.

Relevancia: Juzgue si cada ítem en particular es importante y apropiado para incluirlo en el cuestionario, considerando el marco. Opine sobre la relevancia de cada ítem usando los términos “Relevante”, “Con dudas” o “No relevante”.

Ítems tipo I: Concepto de población.

1.- ¿Qué entiendes por una población en estadística?

(Respuesta abierta)

2.- En una escuela se realiza un concurso de tiros libres a la canasta de basquetbol. Cada estudiante lanza 10 tiros y se registra el número de éxitos. ¿Cuál de las siguientes opciones elegirías como la población del experimento anterior?

- a) Los estudiantes de dicha escuela.
- b) El conjunto $I = \{0, 1\}$, donde 1 = “Encestar”, 0 = “No encestar”.
- a) El conjunto de todas las secuencias o arreglos de 0 y 1 de tamaño 10. Ejemplo: (1,0,1,0,1,0,0,0,1,0,1).
- c) Los valores de la variable “El número de tiros anotados en 10 intentos”. Es decir, $A = \{0, 1, 2, \dots, 10\}$ y su distribución.
- d) Otra. Explica tu respuesta.

Por favor, explica a detalle por qué elegiste la opción de la pregunta anterior.

(Respuesta abierta)

3.- Se saca al azar una ficha de una urna de 500 fichas blancas y 800 fichas negras y se observa el color de ficha extraída, ¿Cuál de las siguientes opciones elegirías para ser la población relativa a la extracción de una ficha?

- a) Las 1300 fichas (500 blancas y 800 negras).
- b) El conjunto $\{B, N\}$ donde B es blanca y N negra.
- c) La distribución de Bernoulli con $B = “1”$, $N = “0”$, $P(1)=500/1300$ y $P(0) =800/130$.
- d) Otra. Explica tu respuesta en la siguiente pregunta.

Por favor, explica a detalle por qué elegiste la opción de la pregunta anterior.
(respuesta abierta)

Evaluación de los ítems. Califique la formulación, pertinencia y relevancia de cada ítem en relación con el concepto en cuestión. Marque con una “X” en la casilla que considere según su criterio.

No.	CONCEPTOS / ÍTEMS	Formulación			Pertinencia			Relevancia			Observaciones / Sugerencias
		Adecuada	A mejorar	No Adecuada	Pertinente	Con dudas	No Pertinente	Relevante	Con dudas	No Relevante	
	Concepto 1: Población										
1	¿Qué entiendes por una población en estadística?										
2	Cada estudiante lanza 10 tiros y se registra el número de éxitos. ¿Cuál de las siguientes opciones elegirías como la población del experimento anterior?										
3	¿Cuál de las siguientes opciones elegirías para ser la población relativa a la extracción de una ficha?										

Ítems tipo 2. Concepto de Muestra

4.- ¿Qué entiendes por una muestra en estadística?
(respuesta abierta)

5.- En el mismo contexto del problema 1 sobre los tiros libres en el basquetbol; se quiere obtener una muestra de tamaño 5. ¿Cuál de los siguientes incisos elegirías como muestra?

- El número de anotaciones de 5 estudiantes después de que cada quién lance 10 tiros.
- Cinco secuencias o arreglos de ceros (“0”) y unos (“1”) de tamaño 10.
- Cinco valores con posibles repeticiones del conjunto $A = \{0, 1, 2, \dots, 10\}$.
- Cinco valores aleatorios de la variable $X =$ “El número de tiros anotados en 10 intentos” de acuerdo con su distribución.
- Otra. Explica tu respuesta en la siguiente pregunta.

Por favor explica a detalle por qué elegiste la opción de la pregunta anterior.
(Respuesta abierta)

6.- Se quiere obtener una muestra de tamaño 10 de una urna con 500 fichas blancas y 800 fichas negras ¿Cuál de los siguientes incisos elegirías como muestra?

- a) Diez fichas sacadas de la urna.
- b) Una secuencia o arreglo de tamaño 10 donde cada entrada es una "B" o "N".
- c) Diez valores aleatorios de la distribución de Bernoulli con "B" $=1$, "N" $=0$, $P(1)=500/1300$ y $P(0)=800/1300$.
- d) Otra. Explica tu respuesta en la siguiente pregunta.

Por favor explica a detalle por qué elegiste la opción de la pregunta anterior.

(Respuesta abierta)

Evaluación de los ítems. Califique la formulación, pertinencia y relevancia de cada ítem en relación con el concepto en cuestión. Marque con una "X" en la casilla que considere según su criterio.

No.	CONCEPTOS / ÍTEMS	Formulación			Pertinencia			Relevancia			Observaciones / Sugerencias
		Adecuada	A mejorar	No Adecuada	Pertinente	Con dudas	No	Relevante	Con dudas	No Relevante	
4	¿Qué entiendes por una muestra en estadística?										
5	En el mismo contexto del problema 1 sobre los tiros libres en el basketbol; se quiere obtener una muestra de tamaño 5. ¿Cuál de los siguientes incisos elegirías como muestra?										
6	Se quiere obtener una muestra de tamaño 10 de una urna con 500 fichas blancas y 800 fichas negras ¿Cuál de los siguientes incisos elegirías como muestra?										

Ítems tipo III: Muestreo repetido

7.- ¿Qué entiendes por un muestreo repetido?

(Respuesta abierta)

8.- Se quiere realizar un muestreo repetido de tamaño 100 en el contexto de los tiros a la canasta de basketbol. ¿Cuál de las siguientes opciones elegirías para hacerlo?

- f) Tomar a un alumno al azar de la escuela, se registra el número de anotaciones en 10 tiros. El alumno elegido ya no puede ser seleccionado nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- g) Tomar a un alumno al azar de la escuela, se registra el número de anotaciones en 10 tiros. El alumno elegido podría ser seleccionado nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- h) Tomar 10 alumnos al azar de la escuela, se registra el número de anotaciones en 10 tiros de cada uno. Los 10 alumnos elegidos no pueden ser seleccionados nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- i) De la lista de todos los alumnos se eligen 10 alumnos al azar de la escuela, se registra el número de anotaciones en 10 tiros de cada uno. Se toma otra muestra de la misma manera (algunos podrían volver a caer en la muestra). Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- j) Otro procedimiento. Explica tu respuesta.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior
(Respuesta abierta)

El siguiente problema es una versión modificada del ejercicio que aparece en la quinta sección de Chance et al. (2004).

9.- Se tiene una caja con fichas negras y blancas, pero se desconoce la cantidad exacta de cada color, solo se sabe que suman 2000 fichas.

¿Cuál crees que sería la mejor forma de estimar la proporción de fichas negras si se te permite sacar 20 muestras de 10 fichas?

- a) Sacar una muestra aleatoria, registrar la proporción de fichas negras, dejar fuera esa muestra de la caja y volver a sacar una muestra nueva, así sucesivamente. Al final observar cuál proporción apareció más veces.
- b) Sacar una muestra aleatoria, registrar la proporción de fichas negras, dejar fuera esa muestra de la caja y volver a sacar una muestra nueva, así sucesivamente. Al final, sumar la proporción de todas las muestras obtenidas y dividir entre 20.
- c) Sacar una muestra aleatoria, registrar la proporción de fichas negras, regresar la muestra a la caja y volver a sacar una muestra nueva, así sucesivamente. Al final observar cuál proporción apareció más veces.
- d) Sacar una muestra, registrar la proporción de fichas negras, regresar la muestra a la caja y volver a sacar una muestra nueva, así sucesivamente. Al final, sumar la proporción de todas las muestras obtenidas y dividir entre 20.
- e) Otra. Explica tu respuesta abajo.

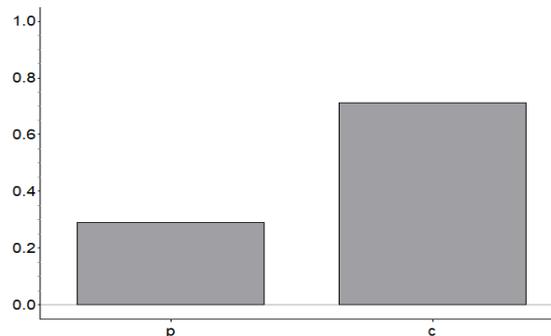
Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior
(Respuesta abierta)

Evaluación de los ítems. Califique la formulación, pertinencia y relevancia de cada ítem en relación con el concepto en cuestión. Marque con una “X” en la casilla que considere según su criterio.

N°	CONCEPTOS / ÍTEMS	Formulación			Pertinencia			Relevancia			Observaciones / Sugerencias
		Adecuada	A mejorar	No Adecuada	Pertinente	Con dudas	No Pertinente	Relevante	Con dudas	No Relevante	
7	¿Qué entiendes por un muestreo repetido?										
8	Se quiere realizar un muestreo repetido de tamaño 100 en el contexto de los tiros a la canasta de basquetbol. ¿Cuál de las siguientes opciones elegirías para hacerlo?										
9	¿Cuál crees que sería la mejor forma de estimar la proporción de fichas negras si se permite sacar 20 muestras de 10 fichas?										

Ítems tipo 4. Distribución muestral

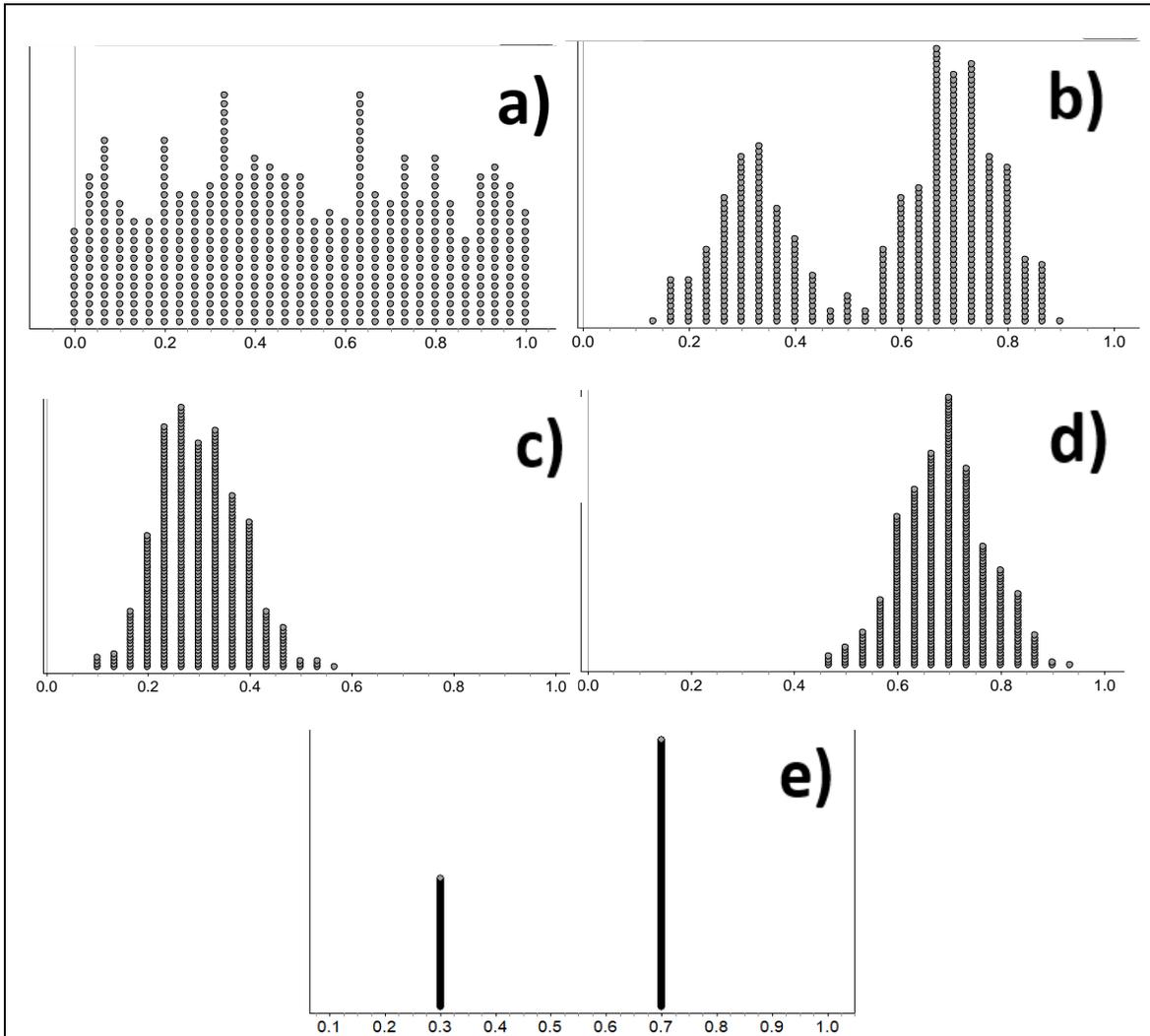
10.- La siguiente imagen corresponde a la distribución de la probabilidad de que un alumno anote o no un tiro libre. Se encesta con una probabilidad de 0.3 y la probabilidad de no encestar es de 0.7.



Se simulan 500 muestras de tamaño 30 mediante una computadora, cada muestra simula el lanzamiento de 30 tiros libres a la canasta con base en la distribución de probabilidad establecida previamente. De cada muestra se registra la proporción de encestes, es decir, se cuenta el número de veces que el tiro es acertado y se divide entre 30, obteniendo un número entre 0 y 1.

La proporción de cada muestra se representa mediante un punto en una gráfica, y si varias proporciones son iguales, sus puntos en la gráfica se apilan.

¿Cuál de las siguientes gráficas representa la distribución de las proporciones de los encastes de las 500 muestras simuladas?



¿Por qué elegiste la(s) opción(es) en la pregunta anterior?

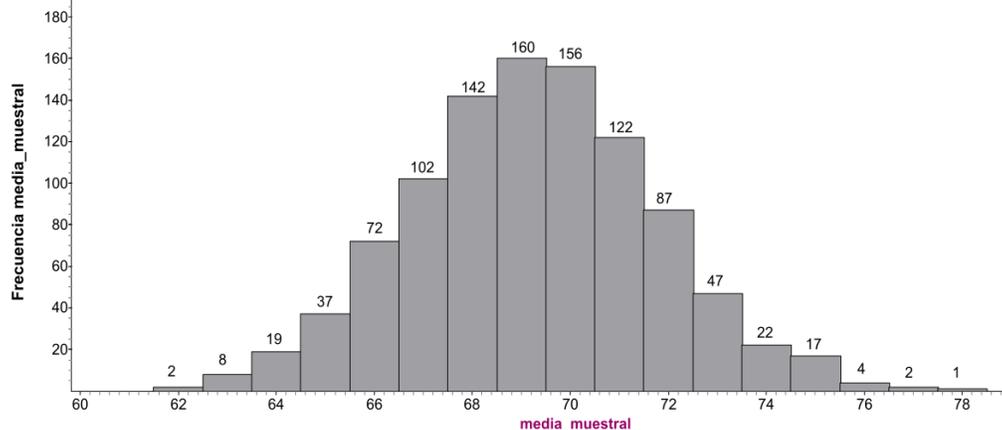
(Respuesta abierta)

El siguiente problema es una versión modificada del problema que aparece en Batanero et al. (2018)

11.- Se sabe que la esperanza de vida de los países se distribuye de forma normal con una media 69.2 años y desviación estándar de 15. A partir de los datos mundiales se simula una muestra aleatoria de tamaño 35 y se obtiene su media. Este procedimiento se repite 1000

veces y se obtiene la distribución de frecuencias de las medias de las 1000 muestras obtenidas en la simulación.

La siguiente gráfica muestra la distribución correspondiente. Cada número sobre las barras es su frecuencia absoluta.



Si se obtuviera una muestra real de países de Europa de tamaño 35 y se calculara su media, ¿aproximadamente a partir de qué valor del estadístico (en este caso la media) de la muestra real es significativo con un nivel de significación del 5%? Es decir, ¿a partir de qué valor del estadístico se podría rechazar la hipótesis de que los países de Europa tienen una esperanza de vida igual a la media mundial? Usa los valores de la distribución de arriba para responder la pregunta.

- Si el estadístico tiene un valor mayor a 69.2.
- A partir de 74, porque 74 o más tienen aproximadamente una frecuencia del 5% del total.
- Si es exactamente 73, porque tiene un valor cercano al 5% del total.
- Si el estadístico cae en el intervalo de $[65, 74]$, porque los valores fuera de ese intervalo, es decir los extremos de la distribución tienen una frecuencia aproximada del 5% del total.
- Otro valor. explica tu respuesta

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior

(Respuesta abierta)

Evaluación de los ítems. Califique la formulación, pertinencia y relevancia de cada ítem en relación con el concepto en cuestión. Marque con una “X” en la casilla que considere según su criterio.

Nº	CONCEPTOS / ÍTEMS	Formulación			Pertinencia			Relevancia			Observaciones / Sugerencias
		Adecuada	A mejorar	No Adecuada	Pertinente	Con dudas	No Pertinente	Relevante	Con dudas	No Relevante	
	Concepto 4: Distribución muestral Empírica										
10	¿Cuál de las siguientes gráficas representa la distribución de las proporciones de los encuestes de las 500 muestras simuladas?										
11	¿A partir de qué valor del estadístico se podría rechazar la hipótesis de que los países de Europa tienen una esperanza de vida igual a la media mundial?										

Opinión sobre la aplicabilidad del cuestionario:

Aplicable sin modificaciones []

Aplicable después de corregir observaciones menores []

Aplicable después de corregir observaciones mayores []

No aplicable []

Observaciones generales:

Apéndice D: cuestionario aplicado a los modelos de inteligencia artificial

Ítems tipo I: Concepto de población.

1.- ¿Qué entiendes por una población en estadística?

(Respuesta abierta)

2.- En una escuela se realiza un concurso de tiros libres a la canasta de basquetbol. Cada estudiante lanza 10 tiros y se registra el número de tiros encestandos. ¿Cuál de las siguientes opciones elegirías como la población del experimento anterior?

- e) Los estudiantes de dicha escuela.
- f) El conjunto $I = \{0, 1\}$, donde 1 = “Encestar”, 0 = “No encestar”.
- g) El conjunto de todas las secuencias o arreglos de “0” y “1” de tamaño 10, donde 1 = “Encestar”, 0 = “No encestar”. Ejemplo de una secuencia de dicho conjunto:
(1,0,1,0,1,0,0,0,1,0,1).
- h) Los valores de la variable: “El número de tiros anotados en 10 lanzamientos”. Es decir, $A = \{0, 1, 2, \dots, 10\}$ y su distribución.
- i) Otra. Explica tu respuesta.

Por favor, explica a detalle por qué elegiste la opción de la pregunta anterior.

3.- Se saca al azar una sola ficha de una urna de 500 fichas blancas y 800 fichas negras y se observa el color de ficha extraída, ¿Cuál de las siguientes opciones elegirías para ser la población asociada a la extracción de una ficha de la urna?

- e) Las 1300 fichas (500 blancas y 800 negras).
- f) El conjunto $\{B, N\}$ donde B representa una ficha blanca y N una ficha negra.
- g) Los valores de la variable "El resultado de sacar una única ficha de una urna que solo tiene fichas negras y blancas". Es decir, $B = \{0, 1\}$, donde la probabilidad de sacar una ficha blanca (1) es de 500/1300, mientras que la probabilidad de sacar una ficha negra (0) es de 800/1300.
- h) Otra. Explica tu respuesta en la siguiente pregunta.

Por favor, explica a detalle por qué elegiste la opción de la pregunta anterior.

Ítems tipo II. Concepto de Muestra

4.- ¿Qué entiendes por una muestra en estadística?

(Respuesta abierta)

5.- En el contexto del problema 1 sobre los tiros libres en el basquetbol; se quiere obtener una muestra de tamaño 5. ¿Cuál de los siguientes incisos elegirías como muestra?

- f) El número de anotaciones de 5 estudiantes después de que cada uno haya lanzado 10 tiros.
- g) Cinco secuencias o arreglos de tamaño 10 compuestos por ceros (0) y unos (1), donde "0" significa "no encestar" y "1" significa "encestar". Ejemplo de una secuencia: (0, 1, 0, 1, 1, 0, 1, 1, 0, 1).
- h) Cinco valores, que pueden aparecer más de una vez, seleccionados del conjunto $A = \{0, 1, 2, \dots, 10\}$. Cada valor representa el número de encestes logrados por un alumno después de 10 tiros.
- i) Cinco valores aleatorios de la variable $X =$ "El número de tiros anotados en 10 intentos" de acuerdo con su distribución.
- j) Otra. Explica tu respuesta en la siguiente pregunta.

Por favor explica a detalle por qué elegiste la opción de la pregunta anterior.

6.- Se quiere obtener una muestra de tamaño 10 de una urna con 500 fichas blancas y 800 fichas negras ¿Cuál de los siguientes incisos elegirías como muestra?

- e) Diez fichas sacadas de la urna.
- f) Una secuencia o arreglo de tamaño 10 compuestos por ceros (0) y unos (1), donde "0" significa "ficha negra" y "1" significa "ficha blanca". Ejemplo de una secuencia: (0, 1, 0, 1, 1, 0, 1, 1, 0, 1).
- g) Diez valores aleatorios de la variable "El resultado extraer una única ficha de una urna que solo tiene fichas negras y blancas". Es decir, $B = \{0, 1\}$, donde la probabilidad de sacar una ficha blanca (1) es de $500/1300$, mientras que la probabilidad de sacar una ficha negra (0) es de $800/1300$.
- h) Otra. Explica tu respuesta en la siguiente pregunta.

Por favor explica a detalle por qué elegiste la opción de la pregunta anterior.

Ítems tipo III: Muestreo repetido

7.- ¿Qué entiendes por un muestreo repetido?

8.- Se quiere realizar un muestreo repetido de tamaño 100 en el contexto de los tiros a la canasta de basketbol. ¿Cuál de las siguientes opciones elegirías para hacerlo?

- k) Tomar a un alumno al azar de la escuela y registrar el número de anotaciones en 10 tiros. El alumno elegido ya no puede ser seleccionado nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- l) Tomar a un alumno al azar de la escuela, se registra el número de anotaciones en 10 tiros. El alumno elegido podría ser seleccionado nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- m) Tomar 10 alumnos al azar de la escuela, se registra el número de anotaciones en 10 tiros de cada uno. Los 10 alumnos elegidos no pueden ser seleccionados nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- n) De la lista de todos los alumnos se eligen 10 alumnos al azar de la escuela, se registra el número de anotaciones en 10 tiros de cada uno. Se toma otra muestra de la misma manera (algunos podrían volver a caer en la muestra). Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- o) Otro procedimiento. Explica tu respuesta.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior
(Respuesta abierta)

El siguiente problema es una versión modificada del ejercicio que aparece en la quinta sección de Chance et al. (2004).

9.- Se tiene una caja con fichas negras y blancas, pero se desconoce la cantidad exacta de cada color, solo se sabe que suman 2000 fichas.

¿Cuál crees que sería la mejor forma de estimar la proporción de fichas negras si se te permite sacar 20 muestras de 10 fichas?

- f) Sacar una muestra aleatoria, registrar la proporción de fichas negras, dejar fuera esa muestra de la caja y volver a sacar una muestra nueva, así sucesivamente. Al final observar cuál proporción apareció más veces.
- g) Sacar una muestra aleatoria, registrar la proporción de fichas negras, dejar fuera esa muestra de la caja y volver a sacar una muestra nueva, así sucesivamente. Al final, sumar la proporción de todas las muestras obtenidas y dividir entre 20.
- h) Sacar una muestra aleatoria, registrar la proporción de fichas negras, regresar la muestra a la caja y volver a sacar una muestra nueva, así sucesivamente. Al final observar cuál proporción apareció más veces.

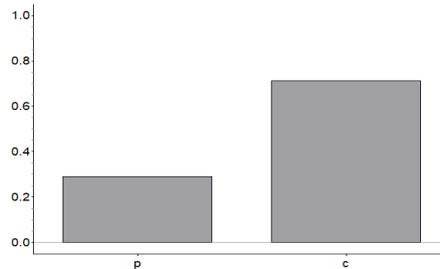
- i) Sacar una muestra, registrar la proporción de fichas negras, regresar la muestra a la caja y volver a sacar una muestra nueva, así sucesivamente. Al final, sumar la proporción de todas las muestras obtenidas y dividir entre 20.
- j) Otra. Explica tu respuesta abajo.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior

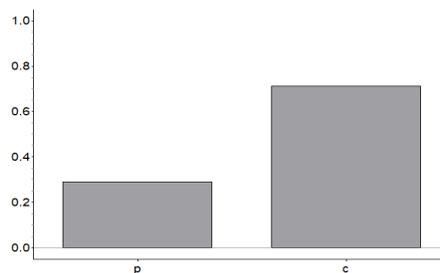
Ítems tipo IV. Distribución muestral

10.- ¿Qué entiendes por una distribución muestral?
(Respuesta abierta)

11.- La siguiente imagen corresponde a la distribución de la probabilidad de que un alumno anote o no un tiro libre. Se encesta con una probabilidad de 0.3 y la probabilidad de no encestar es de 0.7.



11.- La siguiente imagen corresponde a la distribución de la probabilidad de que un alumno anote o no un tiro libre. Se encesta con una probabilidad de 0.3 y la probabilidad de no encestar es de 0.7.



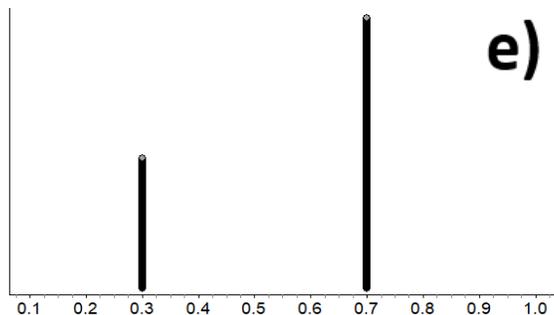
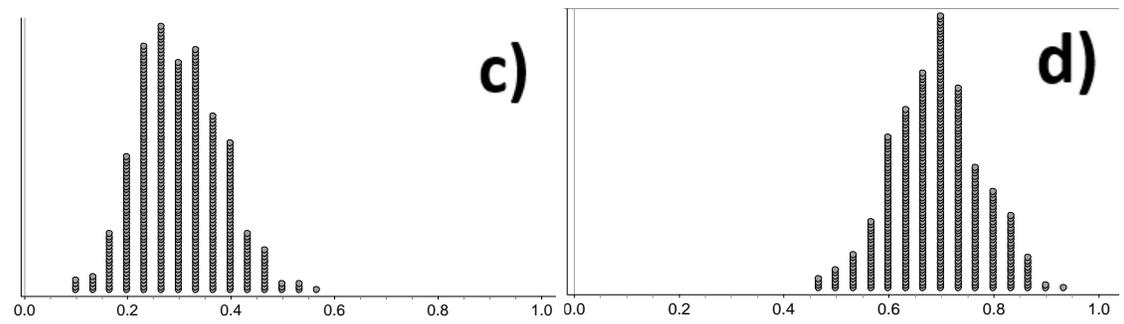
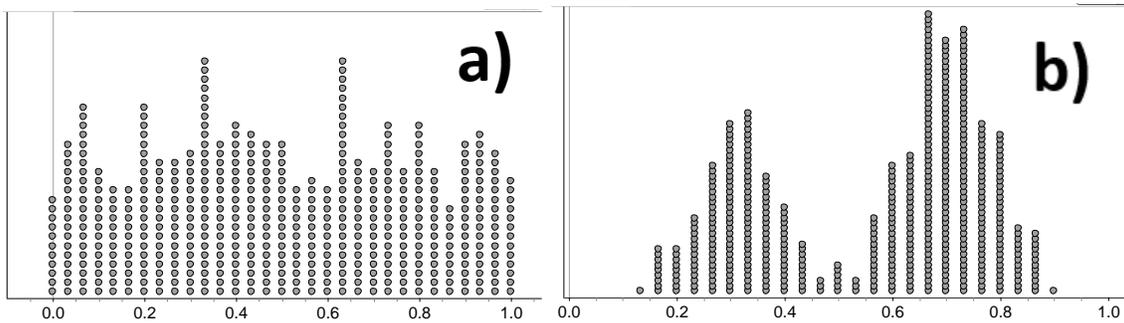
Se simulan 500 muestras de tamaño 30 mediante una computadora, cada muestra simula el lanzamiento de 30 tiros libres a la canasta con base en la distribución de probabilidad establecida previamente. De cada muestra se registra la proporción de encestes, es decir, se cuenta el número de veces que el tiro es acertado y se divide entre 30, obteniendo un número entre 0 y 1.

La proporción de cada muestra se representa mediante un punto en una gráfica, y si varias proporciones son iguales, sus puntos en la gráfica se apilan.

¿Cuál de las siguientes gráficas representa la distribución de las proporciones de los éxitos de las 500 muestras simuladas?

(Verbalización de los gráficos)

- a) Una distribución uniforme de las proporciones de las muestras en el intervalo $[0,1]$.
- b) Una distribución bimodal de las proporciones con modas en 0.3 y 0.7 en el intervalo $[0,1]$.
- c) Una distribución con forma binomial de las muestras con en el intervalo $[0,1]$ con media en 0.3.
- d) Una distribución con forma binomial de las muestras con en el intervalo $[0,1]$ con media en 0.7.
- e) Una distribución con solo dos valores en las proporciones, 0.3 y 0.7 y la mayoría en 0.7.

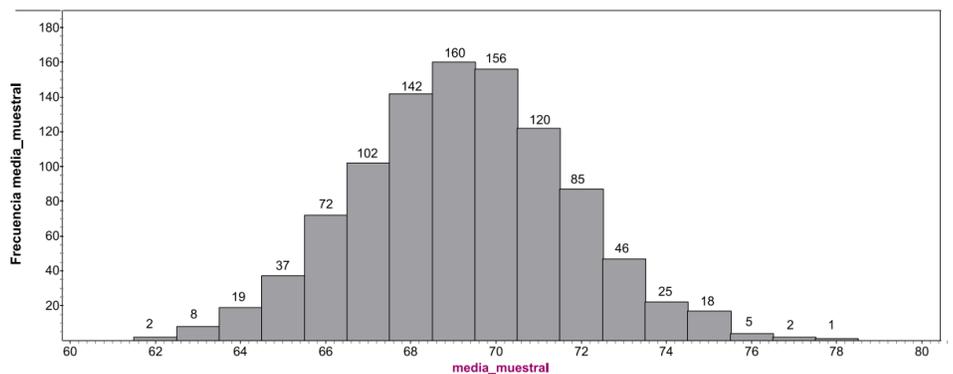


¿Por qué elegiste la(s) opción(es) en la pregunta anterior?

El siguiente problema es una versión modificada del problema que aparece en Batanero et al. (2018)

12.- Se sabe que la esperanza de vida de los países se distribuye de forma normal con una media 69.2 años y desviación estándar de 15. A partir de los datos mundiales se simula una muestra aleatoria de tamaño 35 y se obtiene su media. Este procedimiento se repite 1000 veces y se obtiene la distribución de frecuencias de las medias de las 1000 muestras obtenidas en la simulación.

La siguiente gráfica muestra la distribución correspondiente. Cada número sobre las barras es su frecuencia absoluta.



{(Media muestral, Frecuencia): (62,2), (63,8), (64,19), (65,37), (66,72), (67,102), (68,142), (69,160), (70,156), (71,120), (72,85), (73,46), (74,25), (75,18), (76,5), (77,2), (78,1)}

Si se obtuviera una muestra real de países de Europa de tamaño 35 y se calculara su media, ¿aproximadamente a partir de qué valor del estadístico (en este caso la media) de la muestra real es significativo con un nivel de significación del 5%? Es decir, ¿a partir de qué valor del estadístico se podría rechazar la hipótesis de que los países de Europa tienen una esperanza de vida igual a la media mundial? Usa los valores de la distribución de arriba para responder la pregunta.

- f) Si el estadístico tiene un valor mayor a 69.2.
- g) A partir de 74, porque 74 o más tienen aproximadamente una frecuencia del 5% del total.
- h) Si es exactamente 73, porque tiene un valor cercano al 5% del total.
- i) Si el estadístico cae en el intervalo de [65, 74], porque los valores fuera de ese intervalo, es decir los extremos de la distribución tienen una frecuencia aproximada del 5% del total.
- j) Otro valor. explica tu respuesta

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior

(Respuesta abierta)

Apéndice E: cuestionario aplicado a profesores de CCH

Ítems tipo I: Concepto de población.

<p>1.- ¿Qué entiendes por una población en estadística? (Respuesta abierta)</p>
<p>2.- En una escuela se realiza un concurso de tiros libres a la canasta de basquetbol. Cada estudiante lanza 10 tiros y se registra el número de tiros encestandos. ¿Cuál de las siguientes opciones elegirías como la población del experimento anterior?</p> <p>j) Los estudiantes de dicha escuela.</p> <p>k) El conjunto $I = \{0, 1\}$, donde 1 = “Encestar”, 0 = “No encestar”.</p> <p>l) El conjunto de todas las secuencias o arreglos de “0” y “1” de tamaño 10. Ejemplo de una secuencia: (1,0,1,0,1,0,0,1,0,1).</p> <p>m) Los valores de la variable: “El número de tiros anotados en 10 lanzamientos”. Es decir, $A = \{0, 1, 2, \dots, 10\}$ y su distribución.</p> <p>n) Otra. Explica tu respuesta.</p> <p>Por favor, explica a detalle por qué elegiste la opción de la pregunta anterior. (Respuesta abierta)</p>
<p>3.- Se saca al azar una sola ficha de una urna de 500 fichas blancas y 800 fichas negras y se observa el color de ficha extraída, ¿Cuál de las siguientes opciones elegirías para ser la población asociada a la extracción de una ficha de la urna?</p> <p>i) Las 1300 fichas (500 blancas y 800 negras).</p> <p>j) El conjunto $\{B, N\}$ donde B representa una ficha blanca y N una ficha negra.</p> <p>k) Los valores de la variable "El resultado de sacar una única ficha de una urna que solo tiene fichas negras y blancas". Es decir, $B = \{0, 1\}$, donde la probabilidad de sacar una ficha blanca (1) es de 500/1300, mientras que la probabilidad de sacar una ficha negra (0) es de 800/1300.</p> <p>l) Otra. Explica tu respuesta en la siguiente pregunta.</p> <p>Por favor, explica a detalle por qué elegiste la opción de la pregunta anterior. (respuesta abierta)</p>

Ítems tipo II. Concepto de Muestra

4.- ¿Qué entiendes por una muestra en estadística?

(respuesta abierta)

5.- En el contexto del problema 1 sobre los tiros libres en el basquetbol; se quiere obtener una muestra de tamaño 5. ¿Cuál de los siguientes incisos elegirías como muestra?

- k) El número de anotaciones de 5 estudiantes después de que cada uno haya lanzado 10 tiros.
- l) Cinco secuencias o arreglos de tamaño 10 compuestos por ceros (0) y unos (1), donde "0" significa "no encestar" y "1" significa "encestar". Ejemplo de una secuencia: (0, 1, 0, 1, 1, 0, 1, 1, 0, 1).
- m) Cinco valores, que pueden aparecer más de una vez, seleccionados del conjunto $A = \{0, 1, 2, \dots, 10\}$. Cada valor representa el número de encestes logrados por un alumno después de 10 tiros.
- n) Cinco valores aleatorios de la variable $X = \text{"El número de tiros anotados en 10 intentos"}$ de acuerdo con su distribución.
- o) Otra. Explica tu respuesta en la siguiente pregunta.

Por favor explica a detalle por qué elegiste la opción de la pregunta anterior.

(Respuesta abierta)

6.- Se quiere obtener una muestra de tamaño 10 de una urna con 500 fichas blancas y 800 fichas negras ¿Cuál de los siguientes incisos elegirías como muestra?

- i) Diez fichas sacadas de la urna.
- j) Una secuencia o arreglo de tamaño 10 compuestos por ceros (0) y unos (1), donde "0" significa "ficha negra" y "1" significa "ficha blanca". Ejemplo de una secuencia: (0, 1, 0, 1, 1, 0, 1, 1, 0, 1).
- k) Diez valores aleatorios de la variable "El resultado extraer una única ficha de una urna que solo tiene fichas negras y blancas". Es decir, $B = \{0, 1\}$, donde la probabilidad de sacar una ficha blanca (1) es de $500/1300$, mientras que la probabilidad de sacar una ficha negra (0) es de $800/1300$.
- l) Otra. Explica tu respuesta en la siguiente pregunta.

Por favor explica a detalle por qué elegiste la opción de la pregunta anterior.

(Respuesta abierta)

Ítems tipo III: Muestreo repetido

7.- ¿Qué entiendes por un muestreo repetido?

(Respuesta abierta)

8.- Se quiere realizar un muestreo repetido de tamaño 100 en el contexto de los tiros a la canasta de basketbol. ¿Cuál de las siguientes opciones elegirías para hacerlo?

- p) Tomar a un alumno al azar de la escuela y registrar el número de anotaciones en 10 tiros. El alumno elegido ya no puede ser seleccionado nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- q) Tomar a un alumno al azar de la escuela, se registra el número de anotaciones en 10 tiros. El alumno elegido podría ser seleccionado nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- r) Tomar 10 alumnos al azar de la escuela, se registra el número de anotaciones en 10 tiros de cada uno. Los 10 alumnos elegidos no pueden ser seleccionados nuevamente. Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- s) De la lista de todos los alumnos se eligen 10 alumnos al azar de la escuela, se registra el número de anotaciones en 10 tiros de cada uno. Se toma otra muestra de la misma manera (algunos podrían volver a caer en la muestra). Se repite el procedimiento con hasta registrar a los 100 estudiantes.
- t) Otro procedimiento. Explica tu respuesta.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior

(Respuesta abierta)

El siguiente problema es una versión modificada del ejercicio que aparece en la quinta sección de Chance et al. (2004).

9.- Se tiene una caja con fichas negras y blancas, pero se desconoce la cantidad exacta de cada color, solo se sabe que suman 2000 fichas.

¿Cuál crees que sería la mejor forma de estimar la proporción de fichas negras si se te permite sacar 20 muestras de 10 fichas?

- k) Sacar una muestra aleatoria, registrar la proporción de fichas negras, dejar fuera esa muestra de la caja y volver a sacar una muestra nueva, así sucesivamente. Al final observar cuál proporción apareció más veces.
- l) Sacar una muestra aleatoria, registrar la proporción de fichas negras, dejar fuera esa muestra de la caja y volver a sacar una muestra nueva, así sucesivamente. Al final, sumar la proporción de todas las muestras obtenidas y dividir entre 20.

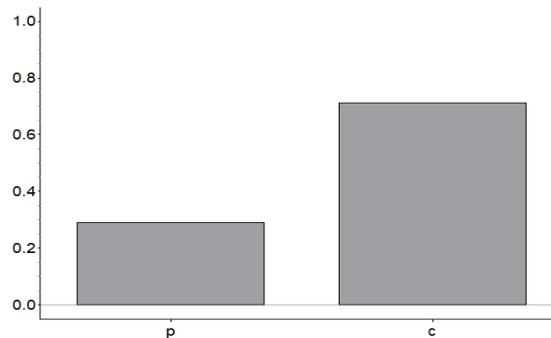
- m) Sacar una muestra aleatoria, registrar la proporción de fichas negras, regresar la muestra a la caja y volver a sacar una muestra nueva, así sucesivamente. Al final observar cuál proporción apareció más veces.
- n) Sacar una muestra, registrar la proporción de fichas negras, regresar la muestra a la caja y volver a sacar una muestra nueva, así sucesivamente. Al final, sumar la proporción de todas las muestras obtenidas y dividir entre 20.
- o) Otra. Explica tu respuesta abajo.

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior
(Respuesta abierta)

Ítems tipo IV. Distribución muestral

10.- ¿Qué entiendes por una distribución muestral?
(Respuesta abierta)

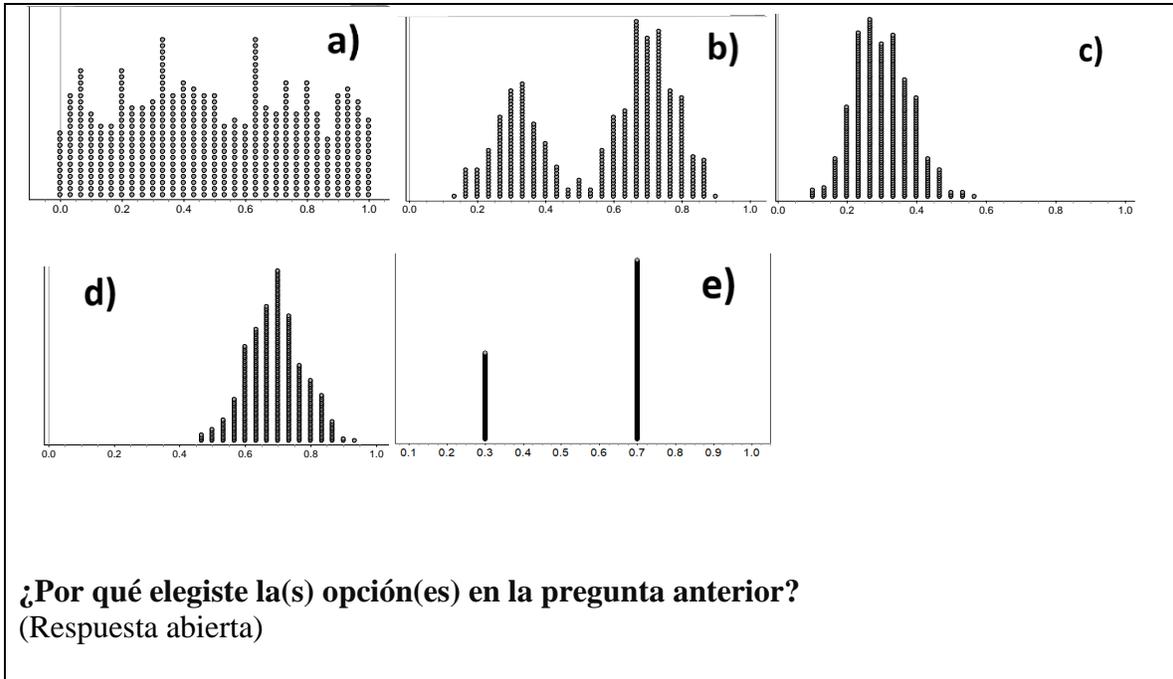
11.- La siguiente imagen corresponde a la distribución de la probabilidad de que un alumno anote o no un tiro libre. Se encesta con una probabilidad de 0.3 y la probabilidad de no encestar es de 0.7.



Se simulan 500 muestras de tamaño 30 mediante una computadora, cada muestra simula el lanzamiento de 30 tiros libres a la canasta con base en la distribución de probabilidad establecida previamente. De cada muestra se registra la proporción de encestes, es decir, se cuenta el número de veces que el tiro es acertado y se divide entre 30, obteniendo un número entre 0 y 1.

La proporción de cada muestra se representa mediante un punto en una gráfica, y si varias proporciones son iguales, sus puntos en la gráfica se apilan.

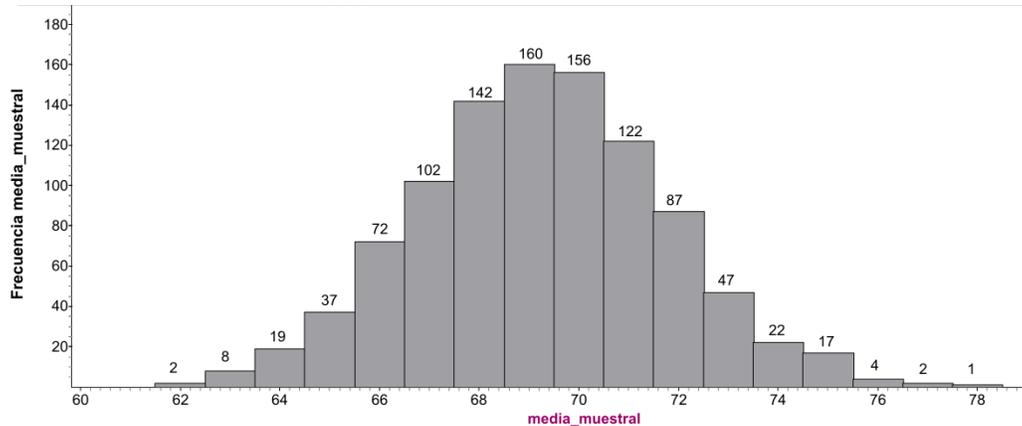
¿Cuál de las siguientes gráficas representa la distribución de las proporciones de los encestes de las 500 muestras simuladas?



El siguiente problema es una versión modificada del problema que aparece en Batanero et al. (2018)

12.- Se sabe que la esperanza de vida de los países se distribuye de forma normal con una media 69.2 años y desviación estándar de 15. A partir de los datos mundiales se simula una muestra aleatoria de tamaño 35 y se obtiene su media. Este procedimiento se repite 1000 veces y se obtiene la distribución de frecuencias de las medias de las 1000 muestras obtenidas en la simulación.

La siguiente gráfica muestra la distribución correspondiente. Cada número sobre las barras es su frecuencia absoluta.



Si se obtuviera una muestra real de países de Europa de tamaño 35 y se calculara su media, ¿aproximadamente a partir de qué valor del estadístico (en este caso la media) de la muestra real es significativo con un nivel de significación del 5%? Es decir, ¿a partir de qué valor del estadístico se podría rechazar la hipótesis de que los países de Europa tienen una esperanza de vida igual a la media mundial? Usa los valores de la distribución de arriba para responder la pregunta.

- k) Si el estadístico tiene un valor mayor a 69.2.
- l) A partir de 74, porque 74 o más tienen aproximadamente una frecuencia del 5% del total.
- m) Si es exactamente 73, porque tiene un valor cercano al 5% del total.
- n) Si el estadístico cae en el intervalo de [65, 74], porque los valores fuera de ese intervalo, es decir los extremos de la distribución tienen una frecuencia aproximada del 5% del total.
- o) Otro valor. explica tu respuesta

Por favor explica a detalle por qué elegiste la(s) opción(es) de la pregunta anterior
(Respuesta abierta)

Apéndice F: modificaciones realizadas en cada fase de implementación

Figura 7.1

Modificaciones realizadas en cada fase de implementación

Fase de implementación	Modificaciones realizadas
Cuatro estudiantes	Posterior a la implementación, para mantener un mismo contexto en los ítems de opción múltiple, cada inciso se encuentra vinculado al mismo experimento y representa alguno de los niveles de razonamiento.
Tres expertos en educación	Se llevaron a cabo modificaciones que se centraron en aspectos de redacción. Estas modificaciones fueron orientadas por las opiniones de los especialistas, quienes señalaron que ciertos fragmentos podrían generar confusión o parecer demasiado abstractos para los estudiantes. Como resultado, se presentan las variables aleatorias con una redacción más accesible y menos formal.
Dos modelos de lenguaje GPT	Después de la implementación, se llevaron a cabo dos ajustes en el ítem 12: en primer lugar, se reformuló la pregunta para que se interpretara como una prueba de cola derecha en lugar de una prueba de dos colas. Además, se modificó la redacción de las opciones presentadas, eliminando las justificaciones de cada inciso y manteniendo solo valores o intervalos numéricos.
Profesores De CCH	Después de la aplicación, se llevó a cabo una modificación en el ítem 2. El propósito de esta modificación fue establecer una probabilidad constante en los lanzamientos de tiro libre, lo que posibilita la comparación precisa del desempeño entre los distintos participantes.
