Uniciencia Vol. 37(1), pp. 1-20, January-December, 2023
www.revistas.una.ac.cr/uniciencia
revistauniciencia@una.cr
DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

Conocimiento didáctico-matemático sobre la distribución de la media muestral de profesorado de bachillerato en formación

Prospective high school mathematics teachers' didactic-mathematical knowledge about the sampling distribution of the mean

Conhecimento didático-matemático sobre a distribuição da média amostral de professores bacharéis estagiários

Silvia M. Valenzuela-Ruiz[1], **Carmen Batanero**[1], **Nuria Begué**[2*], **José A. Garzón-Guerrero**[1]

Received: Apr/3/2022 • Accepted: Aug/22/2022 • Published: Jan/1/2023

Resumen

[Objetivo] El objetivo del trabajo es evaluar el conocimiento didáctico-matemático de profesores españoles en formación sobre la distribución de la media muestral, específicamente, el conocimiento común del contenido, y las facetas epistémicas y cognitiva del conocimiento didáctico. **[Metodología]** Se pidió a una muestra de futuros profesores resolver un problema propuesto a los estudiantes en las pruebas de acceso a la universidad, además de resolverlo debían identificar los conceptos, propiedades y procedimientos requeridos para su resolución y los errores previsibles de los estudiantes en este proceso. **[Resultados]** Los resultados de la evaluación fueron muy buenos en lo que se refiere al conocimiento matemático común del tema, aunque se observaron algunos errores, como la confusión de la distribución de la variable analizada en la población con la distribución muestral del estadístico. Los participantes mostraron un desempeño razonable de análisis de los objetos matemáticos requeridos (conceptos, procedimientos y propiedades) para solucionar la tarea propuesta. Fue menor la competencia de análisis de los posibles errores que podrían cometer los estudiantes en la resolución de la tarea. **[Conclusiones]** El estudio revela puntos de mejora en la formación de los futuros profesores sobre la distribución de la media muestral, un contenido relevante para la posterior comprensión del resto de la inferencia. Dicha formación debiera enfatizar la diferencia entre las tres distribuciones de probabilidad que aparecen en el muestreo y la diferencia entre estadístico y parámetro, pues el futuro profesorado no reconoce la posibilidad de errores de este tipo en su alumnado.

Palabras clave: Distribución de la media muestral; conocimiento didáctico-matemático; profesores de bachillerato en formación; evaluación.

* Autor para correspondencia
Silvia M. Valenzuela-Ruiz, svalenzuela@ugr.es, https://orcid.org/0000-0001-7467-8672
Carmen Batanero, batanero@ugr.es, https://orcid.org/0000-0002-4189-7139
Nuria Begué, nbegue@unizar.es, https://orcid.org/0000-0003-1369-8711
José A. Garzón-Guerrero, jgarzon@ugr.es, https://orcid.org/0000-0002-9397-3439
1 Departamento de Didáctica de la Matemática, Universidad de Granada, Granada, España.
2 Departamento de Matemáticas, Área de Didáctica de la Matemática, Universidad de Zaragoza, Zaragoza, España.

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

Abstract

[Objective] The aim of this study is to evaluate the didactic-mathematical knowledge of prospective Spanish teachers about the sampling distribution of the mean – specifically, basic knowledge about content, as well as epistemic and cognitive aspects of didactic knowledge. **[Methodology]** A sample of prospective teachers were asked to solve a problem presented to students in the university's entrance exams, to identify the concepts, properties and procedures required for its solution, and to describe the foreseeable errors of the students in this process. **[Results]** The results obtained showed very good levels of common mathematical knowledge, although some errors were observed, such as confusing population distributions with sampling distributions. The participants were reasonably competent in analyzing the mathematical objects (concepts, procedures and properties) required to solve the proposed task, but the level of competence in identifying possible student errors in solving the task was lower. **[Conclusions]** The study identifies areas for improvement in the training of prospective teachers about the sampling distribution of the mean, which should be well understood when making inferences. Such training should emphasize the difference between the three probability sampling distributions and the difference between statistics and parameters, since prospective teachers do not recognize the possibility of this type of error by their students.

Keywords: distribution of the sample mean; didactic-mathematical knowledge; prospective high-school teachers; evaluation.

Resumo

[**Objetivo**] O objetivo do trabalho é avaliar o conhecimento didático-matemático dos professores estagiários espanhóis sobre a distribuição da média amostral, especificamente, o conhecimento comum do conteúdo, e as facetas epistêmicas e cognitivas do conhecimento didático. [**Metodologia**] Foi solicitado a uma amostra de futuros professores resolver um problema proposto aos alunos nas provas de vestibular e, além de resolvê-lo, eles tiveram que identificar os conceitos, as propriedades e os procedimentos necessários para sua resolução e os erros previsíveis dos alunos neste processo.[**Resultados**] Os resultados da avaliação foram muito bons em termos do conhecimento matemático comum do assunto, embora alguns erros tenham sido observados, tais como a confusão da distribuição da variável analisada na população com a distribuição amostral da estatística. Os participantes mostraram um desempenho razoável na análise dos objetos matemáticos necessários (conceitos, procedimentos e propriedades) para resolver a tarefa proposta. Havia menos competência na análise dos possíveis erros que os estudantes poderiam cometer para resolver a tarefa. **[Conclusões]** O estudo revela pontos de melhoria na capacitação de futuros professores sobre a distribuição da média amostral, um conteúdo relevante para o entendimento subsequente do resto da inferência. Essa capacitação deve enfatizar a diferença entre as três distribuições de probabilidade que aparecem na amostragem e a diferença entre estatístico e parâmetro, pois os futuros professores não reconhecem a possibilidade de tais erros em seus alunos.

Palavras-chave: Distribuição da média amostral; conhecimento didático-matemático; professores bacharéis estagiários; avaliação.

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

Introducción

La inferencia estadística es uno de los temas incluidos actualmente en España en el Bachillerato de Ciencias Sociales, con la finalidad de proporcionar a los estudiantes una base para sus estudios universitarios, a la vez que ampliar su cultura estadística (Gal & Geiger, 2022; Weiland, 2017). Además, es un tema en el que se han descrito numerosos errores cometidos por los estudiantes (Begué et al., 2019; Kula y Koçer, 2020; Makar y Rubin, 2018), por lo que, para asegurar su correcta enseñanza, será necesaria una adecuada preparación de los profesores responsables.

Aunque la investigación sobre el conocimiento del profesor de matemáticas es muy amplia, encontramos pocos referentes que se centren específicamente en lo relacionado con la inferencia estadística. Para proporcionar información al respecto, en este trabajo se evalúa el conocimiento didáctico-matemático (Godino, 2009; Godino et al., 2017; Pino-Fan y Godino, 2015) sobre la distribución de la media muestral, que es la base para el posterior estudio del contraste de hipótesis e intervalo de confianza, temas todos ellos incluidos en el currículo pasado y actual del Bachillerato de Ciencias Sociales en España (MECD, 2015; MEFP, 2022). Para lograr nuestro objetivo, se analizan las respuestas a una tarea abierta en una muestra de 61 estudiantes que se preparan como profesores de matemáticas en España en un programa de Máster que es obligatorio para concursar a una plaza docente en este nivel educativo.

Seguidamente se describen el marco teórico que sustenta la investigación y la metodología utilizada en el trabajo. Se continúa presentando los resultados, finalizando con algunas conclusiones.

Marco teórico

El trabajo se apoya en los diferentes tipos de distribución que se pueden encontrar en el muestreo, el modelo CDM (conocimiento didáctico-matemático), que describe los conocimientos del profesor de matemáticas en el enfoque ontosemiótico y las investigaciones antecedentes sobre el tema.

Distribuciones implícitas en el estudio del muestreo

En el estudio del muestreo es necesario tener en cuenta y relacionar tres tipos de distribuciones de probabilidad (Harradine et al., 2011):

- La distribución de la variable de interés en la población, en el presente estudio, la distribución normal $N\,(\mu, \sigma)$, caracterizada por su media μ y desviación típica σ.
- La distribución de la misma variable en una muestra aleatoria formada por elementos independientes de dicha población.
- La distribución muestral del estadístico, al considerar muestras repetidas del mismo tamaño. En este caso, la media muestral sigue una distribución normal. $N\,(\mu, \frac{\sigma}{\sqrt{n}})$.

Sin embargo, son muchos los estudiantes que confunden estos tipos de distribuciones o no llegan a diferenciar el estadístico (media o varianza en una muestra) y el parámetro (media o varianza en la población) (Begué et al., 2019) y usan los estadísticos en lugar de los parámetros al resolver problemas de inferencia (Espinel et al., 2007; Vallecillos, 1999).

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

Conocimientos del profesor de matemáticas

El conocimiento del profesor y su formación es una de las líneas de investigación más amplias y productivas en educación matemática, como muestran, por ejemplo, los trabajos de Beswick y Goos (2018), Llinares (2018) y Scheiner et al. (2019). Dichos estudios proponen diferentes modelos para describir este conocimiento, que tiene diferentes componentes.

En este trabajo se utiliza el modelo CDM (conocimiento didáctico-matemático) (Godino, 2009; Godino et al., 2017; Pino-Fan y Godino, 2015; Pino-Fan et al., 2015; Scheiner, et al., 2019). El CDM se ha originado dentro del enfoque ontosemiótico del conocimiento y la instrucción matemática (EOS) (Godino et al., 2007; 2019), que parte de las ideas de *situación problema* y *práctica matemática* (actuaciones realizadas para resolver estos problemas). El *objeto matemático* se concibe de una forma muy amplia, diferenciándose, además de las situaciones-problemas: a) el lenguaje matemático (términos, símbolos y gráficos, utilizados para representar los objetos matemáticos y operar con ellos); b) los conceptos (generalizaciones que se delimitan por medio de su definición), c) propiedades que relacionan entre sí objetos matemáticos, d) procedimientos, es decir, estrategias o algoritmos; y e) argumentos, utilizados para justificar las propiedades o soluciones de los problemas. El *significado (personal o institucional)* de un objeto matemático se concibe como el conjunto de prácticas matemáticas personales o institucionales asociadas al campo de problemas relacionado con dicho objeto.

En el modelo CDM se consideran tres dimensiones del conocimiento del profesor: matemática; didáctica y meta didáctica. El *conocimiento meta didáctico* del profesor no se considera en este trabajo y es el necesario para llevar a cabo la identificación y gestión de normas y aspectos que regulan la enseñanza y aprendizaje de las matemáticas (Pino-Fan y Godino, 2015).

La dimensión matemática incluye el *conocimiento matemático común*, y el *conocimiento ampliado del contenido* matemático. El conocimiento matemático común es el pertinente para el nivel educativo donde se lleva a cabo la enseñanza del tema, es decir, el que debe enseñar a sus estudiantes. En este estudio sería el conocimiento de la distribución de la media muestral, que se enseña en el Bachillerato de Ciencias Sociales español. El conocimiento ampliado del contenido matemático es más avanzado y permite articular la enseñanza con los niveles superiores (Godino et al., 2017). Incluye lo que se debe enseñar sobre el tema en niveles superiores (en este caso en la universidad) y aspectos históricos y epistemológicos del tema.

Además, el profesor ha de tener un conocimiento *didáctico-matemático especializado,* que incluye las siguientes facetas (Ver Figura 1):

- *Faceta epistémica*: conocimiento específico del profesor sobre el propio contenido matemático que enseña. Es decir, además del conocimiento común y ampliado, el profesor debe conocer y articular la diversidad de significados de un objeto matemático y ser capaz de resolver una tarea a través de diferentes procedimientos, proporcionar varias justificaciones y reconocer el contenido matemático de las tareas. En este estudio se incluye la capacidad de identificar los objetos matemáticos requeridos para la resolución de una tarea.

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

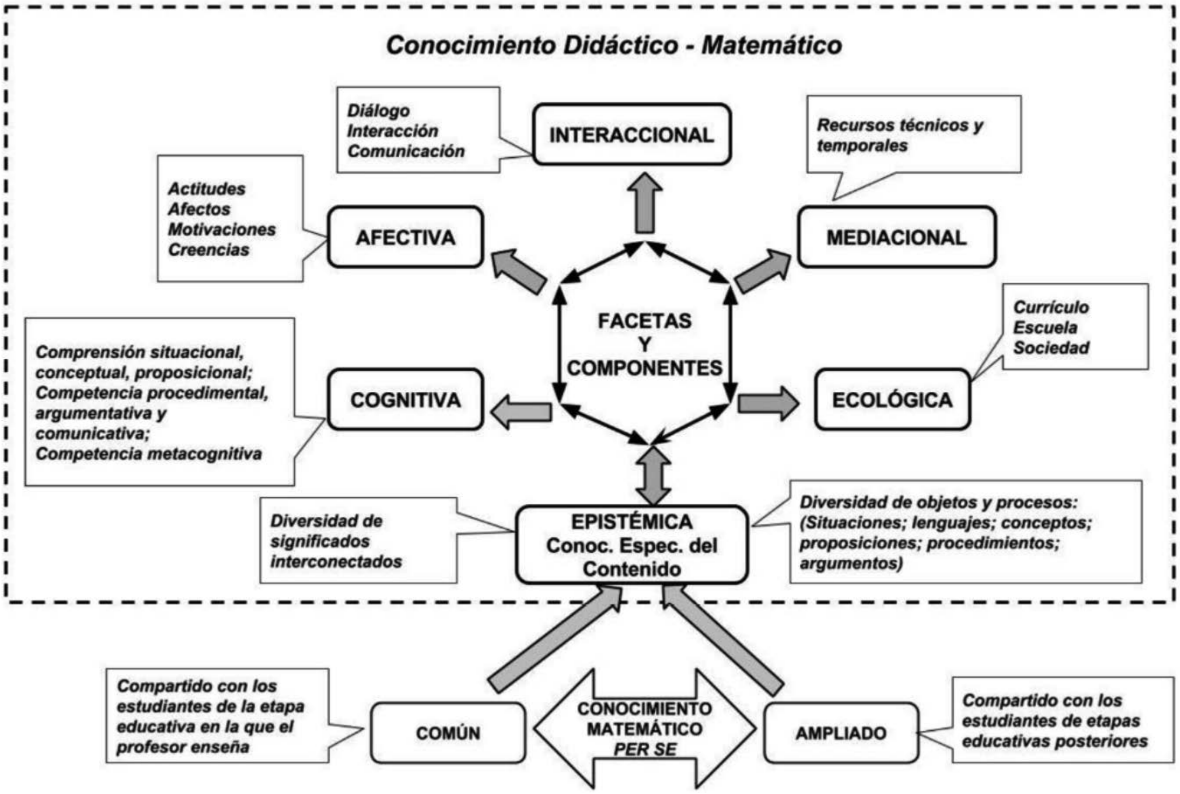

Figura 1. *Facetas y componentes del conocimiento didáctico-matemático.* Extraído de Godino et al. (2017, p. 96).

- *Faceta ecológica:* conocimiento de los aspectos curriculares, contextuales y sociales relacionados con la enseñanza del tema, como el lugar de la distribución muestral en el currículo y su relación con otros contenidos curriculares, o el uso de la inferencia en la investigación, gestión y toma de decisiones.
- *Faceta cognitiva:* conocimiento que tiene el profesor de cómo los estudiantes aprenden, razonan y entienden las matemáticas, de sus estrategias de resolución de problemas y de los errores que cometen en el proceso de aprendizaje, así como sobre su progreso a lo largo del mismo (Pino-Fan y Godino, 2015). Por tanto, implica la competencia para evaluar el grado de aprendizaje de los estudiantes (conocimientos personales), comprender su forma de razonar y reconocer sus dificultades y errores en la resolución de problemas.
- *Faceta afectiva:* conocimientos de las actitudes, emociones, creencias y valores del estudiante en relación con el tema estudiado. Por ejemplo, su gusto o disgusto por la inferencia, si se sienten capaces de aprenderla, o si conceden valor a su estudio para su futura vida profesional.
- *Faceta mediacional*: conocimientos de los recursos didácticos y medios para estudiar el tema, como, por ejemplo, el software disponible para el trabajo de los estudiantes con la inferencia o los libros de texto, apuntes y otros materiales. También incluye el conocimiento

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

del tiempo requerido para la enseñanza y la forma de llevar a cabo su distribución de la mejor forma posible para optimizar el aprendizaje.

- *Faceta interaccional*: capacidad de establecer agrupaciones e interacciones en la clase para permitir identificar y solucionar los problemas que encuentran los estudiantes en el tema.

La evaluación abordada en este trabajo se refiere al conocimiento común sobre la distribución de la media muestral y a las facetas epistémica y cognitiva relacionadas con ella. Es decir, en primer lugar, se evalúa el conocimiento de la distribución de la media muestral que debe enseñar el futuro profesor en el nivel de Bachillerato, mediante la resolución de un problema similar al que debe utilizar con los estudiantes. En segundo lugar se estudian algunos aspectos de su faceta epistémica, para lo cual se propone al futuro profesor que identifique los conceptos, propiedades y procedimientos requeridos para resolver la tarea que se les ha propuesto. Finalmente se evalúa la faceta cognitiva pidiendo a los futuros profesores que identifiquen los posibles errores de sus estudiantes al resolver la tarea propuesta.

Investigaciones previas

La escasa investigación previa sobre el conocimiento del profesor en inferencia se ha centrado, generalmente, en el conocimiento matemático común, es decir, la comprensión que tienen del contenido matemático a enseñar, bien en bachillerato o a nivel universitario. Entre ellas, los trabajos de Batanero et al. (2018), Haller y Krauss (2002), Inzunsa y Jiménez (2013), Liu y Thompson (2009) y Lugo-Armenta y Pino-Fan (2021) estudian el conocimiento común sobre el contraste de hipótesis.

Se han realizado investigaciones que abordan aspectos puntuales de la comprensión del muestreo por parte de futuros profesores, como la de Abu-Ghalyoun (2021), que analiza la comprensión progresiva de la dependencia de la variabilidad muestral respecto al tamaño de la muestra en un futuro profesor, mientras trabaja con la simulación o la de Batanero et al. (2020) en la que se analizan la generación de muestras por parte de futuros profesores. Otras investigaciones con futuros profesores estudian la comprensión de la representatividad y variabilidad (por ejemplo, De Vetten et al., 2018; Watson y Callingham, 2013). Estos dos trabajos se han centrado en futuros profesores de educación primaria y en el contexto de inferencias informales, es decir, sin utilizar los conceptos formales de distribución muestral y sus propiedades, sino únicamente ideas intuitivas de cómo varían los estadísticos muestrales (como la media o proporción) al repetir la toma de muestras.

Respecto a la faceta cognitiva, destacamos el trabajo de López-Martín et al. (2019), quienes investigan el conocimiento de 70 futuros profesores españoles de educación secundaria y bachillerato sobre los errores de sus estudiantes en el contraste de hipótesis e intervalo de confianza, después de haber resuelto ellos mismos un problema de cada uno de estos tipos. Sus resultados muestran cierto conocimiento de los errores frecuentes por parte de los futuros profesores, aunque les falta precisión para describirlos, en general. Además, existe poca consciencia de los errores relacionados con el nivel de significación y p-valor.

En nuestro trabajo nos centramos en la inferencia frecuencial formal, que es la contemplada en la enseñanza de Bachillerato de Ciencias Sociales en España (MECD, 2015; MEFP, 2022) y en algunas facetas del

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

conocimiento didáctico-matemático sobre el tema de futuros profesores de educación secundaria y bachillerato, aportando nuevo conocimiento en un tema que no ha sido investigado previamente.

Metodología

La muestra participante estuvo formada por 61 estudiantes de un programa de máster, que es obligatorio en España para poder ejercer como profesor de matemáticas en la educación secundaria obligatoria (estudiantes de 12 a 15 años) y bachillerato (estudiantes de 16 a 17 años). Este máster se realiza una vez que se ha finalizado un grado sobre una disciplina específica, en el que no se proporcionan conocimientos pedagógicos. El citado máster se orienta a salvar esta carencia en la formación inicial del profesor (Muñiz-Rodríguez et al., 2016) y se estructura en tres módulos: a) uno común para cualquier especialidad de profesorado (lengua, matemáticas, etc.), que incluye cursos de pedagogía, psicología y sociología de la educación; b) otro específico de matemáticas, que desarrolla complementos matemáticos, historia de la matemática, enseñanza y aprendizaje de la matemática e innovación docente en matemáticas y c) el Prácticum, que incluye un periodo de prácticas de enseñanza de matemáticas en instituciones educativas, con la supervisión de un tutor, además de un trabajo fin de máster. La didáctica de la estadística en la Universidad de Granada es un capítulo del curso de innovación docente, y se dedicaron 6 horas al tema en el curso en que se realizó la evaluación.

El 52 % de la muestra participante había cursado un grado universitario de matemáticas y el resto habían finalizado otras carreras de ciencias (estadística (13,1 %), física (3,3 %), química (1,6 %)), arquitectura (3,3 %) o ingeniería (26,2 %)). Se trata de un muestreo dirigido con selección controlada (Mendenhall et al., 2006), que incluyó a todos los participantes en el citado máster el curso 2020-2021 en la Universidad de Granada.

Tarea propuesta y respuestas esperadas

La tarea que se presenta a los estudiantes se enuncia a continuación. Este problema sobre la distribución de la media muestral se ha tomado de las pruebas de acceso a la Universidad para los estudiantes que cursan el segundo año del Bachillerato de Ciencias Sociales. Es decir, se trata de un tipo de problema que los futuros profesores de la muestra deberán proponer o resolver en sus clases. En consecuencia, la resolución del problema evalúa el conocimiento común del contenido por parte de los futuros profesores.

Problema. La altura de los estudiantes de 2º curso de bachillerato de un centro sigue una ley Normal de media 165 cm y desviación típica 10 cm.

1. ¿Qué distribución sigue la altura media de las muestras de tamaño 25?
2. Se elige al azar una muestra de 25 estudiantes y se les mide la altura. ¿Cuál es la probabilidad de que la altura media de esa muestra supere 160 cm?

Para responder la primera parte del problema, se ha de diferenciar la distribución de la población y la distribución muestral del estadístico (en este caso, la media muestral). Además, se debe recordar que si una distribución es normal N con media μ y desviación típica σ, esto es, $N(\mu, \sigma)$, la media muestral sigue una distribución

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

$N\ (\mu, \frac{\sigma}{\sqrt{n}})$ donde *n* es el tamaño de la muestra. Finalmente, se ha de identificar en el enunciado del problema, el valor de la media y desviación típica de la población y el tamaño de la muestra, para deducir que la distribución de la media muestral pedida es normal *N* (165,2).

En el segundo apartado se debe tipificar, para transformar la distribución de la media muestral en una distribución normal estándar, esto es: $P(\bar{x} > 160) = P\left(\frac{\bar{x} - 165}{2} > \frac{160 - 165}{2}\right)$ $= P\ (Z > -2{,}5)$. Utilizando las propiedades de simetría de la distribución normal, obtiene $P\ (Z > -2{,}5) = P\ (Z < 2{,}5)$. Al buscar en la tabla de la distribución normal N (0,1), se obtiene el valor 0,9938.

En una segunda sesión, a una parte de los participantes (45 futuros profesores) se les plantearon dos tareas adicionales: a) Analizar los conceptos, propiedades y procedimientos que utilizaron para resolver el problema propuesto; y b) Describir los errores que podrían cometer sus futuros estudiantes en la resolución del problema.

La pregunta a) analiza la competencia de análisis epistémico de los futuros profesores y, por tanto, evalúa un componente de la faceta epistémica de su conocimiento. Se espera que identifiquen los siguientes objetos matemáticos:

- *Conceptos*: Variable estadística y variable aleatoria, sus medidas de posición central y dispersión (dos medias y dos desviaciones típicas), distribución de la variable en la población, distribución de la media muestral. Muestra, tamaño de la muestra. Distribución normal, parámetros. Probabilidad. En total 13 conceptos.
- *Propiedades:* Aleatoriedad de la muestra, simetría de la función de densidad de la distribución normal. Teorema central del límite. Relación entre la desviación típica muestral y poblacional. Relación entre la media muestral y poblacional. El área total bajo la curva normal es igual a la unidad. Probabilidad del suceso complementario, es decir, 7 propiedades.
- *Procedimientos:* Cálculo de la desviación típica (o varianza) de la media muestral. Tipificación y cambio de variable. Cálculo de probabilidades utilizando la distribución normal. Lectura de las tablas de la distribución normal. Cinco procedimientos.

Por otro lado, la pregunta b) evalúa parte de la faceta cognitiva, pues se refiere al conocimiento de las posibles dificultades de los estudiantes en el tema. Un análisis a priori de los errores previsibles revela los siguientes:

- *Errores en la interpretación del enunciado e identificación de los datos requeridos.* En primer lugar, hay que leer el enunciado del problema para identificar la media y desviación típica de la población y el tipo de distribución de la altura de los estudiantes (Normal). Además, hay que comprender que en la primera parte se pide la distribución de la media muestral en una muestra de 25 elementos. Y en la segunda, el cálculo de la probabilidad pedida. Se pueden producir errores en cada uno de estos pasos (5 posibles errores).
- *Errores de tipo conceptual.* A lo largo del desarrollo de la tarea, se pueden confundir los tipos de distribuciones

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

implícitos en el proceso de muestreo, esto es la distribución de la población con la de la media muestral (Harradine et al., 2011). Una vez identificadas estas distribuciones, es posible confundir el estadístico (media o desviación típica muestral) con el parámetro (media o desviación típica poblacional) (Espinel et al., 2007; Vallecillos, 1999). También es posible no deducir que la distribución de la media muestral es normal o confundir sus parámetros (Batanero et al., 2004) (4 errores).

- *Errores de tipo procedimental.* En la segunda parte del problema el estudiantado podría aplicar mecánicamente la tipificación o tener problemas al trabajar con desigualdades (González y Ojeda, 2017). Se podría encontrar dificultad al leer la tabla de la distribución normal N (0,1), puesto que hay que interpolar para hallar el valor pedido. También se pueden cometer errores al interpretar la simetría de la distribución normal, que es necesario utilizar para calcular la probabilidad pedida (Batanero et al., 2004; Onyancha y Ogbonnaya, 2022). Finalmente, se podrían realizar errores de tipo aritméticos al completar los cálculos (6 errores en total).

Toda la actividad formó parte de un taller sobre inferencia estadística y su didáctica con dos sesiones de dos horas cada una de duración. Posteriormente a la recogida de las soluciones a las tareas, se corrigieron y discutieron colectivamente y se trabajaron algunos recursos didácticos basados en la simulación para la enseñanza del tema. Todo ello con la finalidad de desarrollar el conocimiento didáctico-matemático de los participantes sobre la inferencia estadística.

Análisis y resultados

En los siguientes apartados se analizan los resultados en la resolución del problema, en la identificación de los objetos matemáticos requeridos para su solución y de los posibles errores de los estudiantes en la tarea.

Conocimiento matemático común

En primer lugar, se analizan las soluciones aportadas por los participantes a cada una de las partes del problema presentado a los futuros profesores. Para cada una de las dos preguntas formuladas, se describen las categorías de respuesta, incluyendo para cada una un ejemplo de respuesta de un participante en el estudio.

En la primera pregunta del problema planteado se pide determinar la distribución de la media muestral de una muestra aleatoria de 25 elementos, a partir de los datos del enunciado sobre la distribución de la variable altura de los estudiantes. Las respuestas a este apartado se han clasificado de la siguiente forma:

R1. Solución correcta. El estudiante identifica los datos, diferenciando la distribución de la población y la distribución de la media muestral. Recuerda que en una distribución normal (μ, σ), la media muestral sigue una distribución $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$, por lo que deduce que la distribución de la media muestral pedida es normal N (165,2). Un ejemplo de respuesta correcta se reproduce a continuación.

> E1: Si la altura de los chicos de la clase de 2º curso de bachillerato sigue una distribución normal, entonces la altura media de las muestras es también

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

normal $N\ (\mu, \frac{\sigma}{\sqrt{n}})$ para todo *n*. Por lo que; $X \dashrightarrow N(165{,}10)\ \bar{X} \dashrightarrow N(165, \frac{10}{\sqrt{25}})$ $= N(165{,}2)$.

R2. Respuesta correcta, donde se han seguido todos los pasos anteriores, pero expresando la distribución normal en función de la varianza, y no de la desviación típica.

> E2: Altura de los estudiantes $N\ (\mu, \sigma^2)$; $\mu = 165$; $\sigma^2 = 100$. Sea $(X_1, X_2, \dots X_n)$ una muestra aleatoria. Para las muestras de tamaño 25 la media muestral $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ tiene aproximadamente una distribución normal con media μ y varianza $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$; $N\ (165{,}4)$.

R3. Respuesta incorrecta, sugiere que no se puede deducir la distribución de la media muestral, porque no se aplica el teorema central del límite al tener la muestra menos de 30 elementos. Se interpreta literalmente una regla de aplicación del teorema central del límite (consistente en que es aplicable en muestras de 30 elementos o más). La respuesta es incorrecta, porque el modelo normal se aplica a la distribución de la media muestral, independientemente del tamaño de la muestra, al ser la población de partida normal:

> E3: El teorema central del límite nos garantiza que si X es una variable aleatoria de media y desviación típica finitas, entonces para n > 30 la media muestral se aproxima mediante una distribución normal. En este caso, no llegamos al corte de n = 30, y no sabemos el número de elementos de la población. Sería conveniente tomar algún dato más.

R4. Respuesta incorrecta, asignando a la distribución muestral una distribución normal tipificada. Ello es incorrecto pues implicaría que la distribución de la media muestral sería la misma para cualquier tamaño de muestra, media y desviación típica.

> E4: Como la altura sigue una distribución normal con desviación típica = 10 cm, la altura media muestral sigue una distribución normal estándar o tipificada de media 0 y desviación típica 1, $N\ (0{,}1)$.

Los resultados en este apartado se presentan en la Tabla 1. Observamos que prácticamente todos los participantes en el estudio obtienen correctamente la distribución muestral pedida, la mayor parte utilizando como parámetro la desviación típica. Tan solo un participante piensa que no puede deducirse la distribución por una interpretación demasiado estricta del teorema central del límite, que en realidad no es requerido en la solución como se ha discutido anteriormente y otros dos atribuyen una distribución de la media muestral $N\ (0{,}1)$.

Tabla 1. *Distribución de respuestas en el apartado a). Distribución de la media muestral*

Respuesta	Frecuencia	Porcentaje
R1. Correcta	54	88,5
R2. Correcta, usando la varianza	4	6,6
R3. Incorrecta. No se puede usar el TCL	1	1,6
R4. Incorrecta. N(0,1)	2	3,3

Nota: Fuente propia de la investigación.

Las respuestas al segundo apartado del problema, que implica la realización del cálculo de algunas probabilidades utilizando la distribución de la media muestral previamente identificada se han clasificado en la forma siguiente:

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

R1. Solución correcta. Se utiliza la tipificación para transformar la distribución de la media muestral en una distribución normal estándar. Se utilizan correctamente las propiedades de simetría de la distribución normal y que el área total bajo la función de densidad es igual a la unidad, lee la tabla adecuadamente y obtiene la solución pedida.

E5: Para poder calcular probabilidades, tipificamos la variable para poder utilizar la tabla de la normal con el cambio: $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$. Por tanto Z ~ N (0,1). Calculamos $P(\bar{x} > 160) = P\left(\frac{\bar{x}-165}{2} > \frac{160-165}{2}\right) =$ $P\ (Z > -2{,}5) = 0{,}9938$.

R2. Respuesta correcta, sin utilizar la simetría de la distribución normal. El estudiante aplica correctamente la tipificación en la distribución de la media muestral, pero no utiliza la propiedad de simetría de dicha distribución y calcula la probabilidad pedida mediante la probabilidad del complementario:

E6:
$(\bar{X} > 160) = P\left(\frac{\bar{X}-165}{2} > \frac{160-165}{2}\right)$
$= P(Z > -2{,}5) = 1 - P(Z \leq -2{,}5)$.

R3. Solución correcta, con incorrecciones en el uso del lenguaje algebraico. El estudiante realiza correctamente la tipificación de la distribución de la media muestral, y utiliza la propiedad de simetría de esta distribución para obtener la probabilidad pedida. No obstante, tiene incorrecciones en el lenguaje algebraico, pues no utiliza el símbolo μ, habitual para referirse a la media de la población y en su solución no está claro si la notación $\bar{X}$ se refiere a la media muestral o poblacional, pues en el primer paréntesis utiliza $\bar{x}$. En todo caso, las dos probabilidades igualadas no son equivalentes.

E7: $P(\bar{x} > B) = P\ (\frac{\bar{X}-\bar{x}}{\sigma/\sqrt{n}} > \frac{B-\bar{x}}{\sigma/\sqrt{n}}) = 0{,}99$

R4. Respuesta incorrecta. El participante confunde la distribución de la altura de los estudiantes en la población y la distribución de la media muestral en muestras de 25 elementos, el cual es un error frecuente entre los estudiantes (Harradine et al., 2011). En consecuencia, para calcular la probabilidad pedida, usa la distribución original de la población.

E8: $X \dashrightarrow N\,(165, 10);\ P(\bar{X} > 160)$
$= P\left(\frac{\bar{X}-165}{10} > \frac{160-165}{10}\right)$
$= P\left(Z > \frac{-5}{10}\right) = P\ (Z > -0{,}5)$

R5. Cálculo correcto de la probabilidad pedida, con interpretación errónea de la probabilidad. El futuro profesor obtiene la solución correcta, pero la interpreta como probabilidad del suceso contrario.

E9. Se tiene que tipificar la variable con el siguiente cambio: $z = \frac{X-\mu}{\sigma\sqrt{n}} = \frac{X-165}{2}$, por lo que Z es $N(0{,}1)$. $P(\frac{\bar{X}-165}{2} > \frac{160-165}{2}) = P(Z > -2{,}5) =$ $P(Z < 2{,}5) = 0{,}9938$. Este resultado se ha obtenido mirando la tabla de la normal, por lo que el resultado será que la probabilidad de que la media no supere 160 cm es del 99,38 %.

R6. Planteamiento correcto con lectura incorrecta de la tabla de la N(0,1). El estudiante identifica los datos del problema y tipifica la variable, pero al leer la tabla de

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

la distribución normal N (0,1) comete un error, y llega a una probabilidad errónea.

E10:
$P(\bar{X} > 160) = P(\bar{X} > 160) = 1 - 0{,}993$

Los resultados de este apartado se presentan en la Tabla 2, donde, de nuevo, la mayoría de participantes lo resuelve correctamente. Sin embargo, una pequeña parte comete algún error elemental, con mayor frecuencia asociado a la interpretación incorrecta. En resumen, los participantes en el estudio mostraron buen conocimiento común del contenido (Godino et al., 2017; Pino-Fan y Godino, 2015) salvo los errores citados, los cuales, sin embargo, son importantes al tratarse de futuros profesores que han de enseñar el tema.

Tabla 2. *Distribución de respuestas en el apartado b). Cálculo de probabilidad.*

Respuesta	Frecuencia	Porcentaje
R1. Correcta	48	78,7
R3. Correcta, sin utilizar la simetría de la d. normal	7	11,5
R2. Correcta, con incorrecciones del lenguaje algebraico	1	1,6
R4, Confunde distribución poblacional y muestral	1	1,6
R5. Cálculo correcto e interpretación incorrecta	3	4,9
R6. Lectura incorrecta de la tabla	1	1,6

Nota: Fuente propia de la investigación.

Conocimientos en la faceta epistémica

Para evaluar los conocimientos de los futuros profesores en la faceta epistémica (Godino et al., 2017), se propuso, a una submuestra de 45 futuros profesores, analizar los objetos matemáticos necesarios para realizar la tarea anteriormente descrita.

En la Tabla 3 se presentan algunos resúmenes estadísticos del número de objetos identificados por cada estudiante en la resolución del problema. Como se observa en la tabla, lo más sencillo fue identificar conceptos, con un número medio de 6,6 conceptos identificados por cada futuro profesor o profesora, habiendo todos ellos sido capaces de encontrar al menos uno y encontrando un caso atípico de un participante que fue capaz de citar 18 conceptos diferentes.

E11. Media, desviación típica, varianza, variabilidad de la muestra, distribución (ley) normal, probabilidad, muestra, distribución, al azar, media muestral, intervalo de confianza, quedar por debajo de un valor, error máximo, tamaño mínimo de la muestra, muestreo aleatorio estratificado, muestreo, aleatorio simple y reemplazamiento.

Sin embargo, se cometen algunos errores al considerar objetos matemáticos que no emergen al resolver la situación planteada. Al revisar la respuesta de E11, el estudiante anota: intervalo de confianza, error máximo y tamaño mínimo de muestra,

Tabla 3. *Resúmenes estadísticos del número de objetos identificados por participante.*

	Mínimo	Máximo	Media	D. típica	N. objetos previsto
Número de conceptos	1	18	6,6	3,4	13
Número de propiedades	0	2	0,9	0,8	7
Número de procedimientos	0	7	2,4	1,5	5
Total	3	25	10,0	4,4	23
Errores	0	6	0,6	1,4	0

Nota: Fuente propia de la investigación.

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

los cuales no se utilizan en la tarea. También utiliza con poca precisión la expresión "quedar por debajo de un valor" para referirse a la relación de desigualdad, que no es un concepto, sino una propiedad.

Los conceptos más frecuentemente identificados fueron los relacionados con el muestreo (en total fueron citados 79), distribución muestral (muestra aleatoria, media y varianza muestral, tamaño muestral, 56 referencias), variable estadística (variable, frecuencia, distribución, 51 citas), medidas de posición central (31 casos) o de dispersión de la variable aleatoria (34) y probabilidad (31).

Hubo un total de 14 citas incorrectas al intervalo de confianza, que no es requerido para resolver la actividad propuesta y un participante indica como concepto la variabilidad media, posiblemente es una forma incorrecta de referirse a la desviación típica. Hacemos notar que todos los conceptos previstos en el análisis de la actividad (variable estadística y variable aleatoria, sus medidas de posición central y dispersión, distribución de la variable en la población, distribución muestral de una media de 25 elementos, muestra, tamaño de la muestra, distribución normal, parámetros, probabilidad) fueron citados, excepto el de parámetro. Aunque los futuros profesores identificaron la necesidad de trabajar con la media y desviación típica de la población no utilizaron expresamente la palabra parámetro.

En segundo lugar, destacan por su frecuencia los procedimientos (entre 0 y 7 identificados por participante, con una media de 2,4), es decir, sobre 2 o 3 por persona. Los más frecuentes fueron la tipificación (30 futuros profesores o profesoras), calcular o expresar probabilidades (26 casos), la lectura de la tabla de la distribución normal N (0,1) (20 casos), el cálculo de la media y varianza de la distribución muestral (14 casos) y la probabilidad del suceso contrario (5 citas). Dos participantes aluden al manejo de desigualdades. En definitiva, se citan alguna vez todos los procedimientos previstos en el análisis a priori de la actividad. Respecto a procedimientos incorrectamente identificados, encontramos el de cálculo del intervalo de confianza (7 estudiantes) y el cálculo del tamaño muestral (3 participantes), innecesarios para resolver la tarea. También en casos aislados se alude a la representación de datos, sumatoria y representación en tablas que no son requeridos en la tarea.

Hay muy pocas propiedades reconocidas en la tarea (0 a 2 por participante, no llegan a 1 de media). Entre ellas encontramos el teorema central del límite (16 casos), la simetría de la distribución normal (15 citas) y propiedades algebraicas (11 casos). No se hace referencia a la aleatoriedad de la muestra, ni a la relación entre la desviación típica muestral y poblacional o entre la media muestral y poblacional. Tampoco se alude a que el área total bajo la curva normal debe ser igual a la unidad o la probabilidad del suceso complementario. Por tanto, fue más difícil para estos futuros profesores reconocer propiedades que conceptos o procedimientos. Recordamos, además, que el teorema central del límite no era estrictamente necesario aplicarlo, aunque fue citado por muchos futuros profesores.

Algún estudiante demuestra falta de comprensión de la diferencia entre concepto, procedimiento y propiedad, como se muestra en el siguiente ejemplo: "Error de cálculo en las propiedades" (listado como procedimiento por un estudiante), cuando las propiedades no implican la realización de cálculos. En resumen, los participantes en el estudio mostraron alguna

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

competencia en el análisis de objetos matemáticos requeridos en la tarea, lo que muestra un conocimiento aceptable en la faceta epistémica. Sería mejorable, no obstante, su competencia para identificar los objetos matemáticos utilizados, pues el número medio de objetos citados en cada categoría es pequeño, en relación con el número previsto. En algunos casos se precisa reforzar la diferenciación entre conceptos, propiedades y procedimientos.

Conocimientos en la faceta cognitiva

Para evaluar la faceta cognitiva del conocimiento didáctico-matemático de los futuros profesores (Godino et al., 2017) se propuso a una submuestra de 45 futuros profesores y profesoras describir los errores previsibles del estudiantado de bachillerato al realizar la tarea.

En la Tabla 4 se presentan algunos resúmenes estadísticos del número de errores identificados por cada futuro profesor o profesora. Lo más sencillo fue prever los errores procedimentales del estudiantado, con un número medio de 2,3 errores identificados por cada futuro profesor o profesora en la muestra. Aunque uno de ellos no identificó ninguno, fue un caso atípico, pues el resto citó al menos uno. Respecto a los errores conceptuales que pueden interferir en el desarrollo de la tarea se identifican entre 1 y 3, con una media de 1,9 siendo atípicos los casos en que encuentran 4 o 5 errores.

El total de errores varía de 1 a 8, de modo que todos los participantes han identificado al menos uno, con una media de 4,2 errores descritos por cada sujeto de la muestra, lo cual es muy poco, al comparar con lo esperado en el análisis a priori. Al confrontar con la investigación de López-Martín et al. (2019) en que se pide a los futuros profesores identificar los posibles errores al resolver problemas de contraste de hipótesis e intervalo de confianza, el número medio de errores citados fue similar, unos 7 errores en promedio entre dos problemas, donde el número esperado de posibles errores era también similar (mientras en nuestra investigación solo se ha dado un problema). Reproducimos el ejemplo de un futuro profesor que identifica 8 errores diferentes.

> E12: Respecto a los errores que pueden cometer los estudiantes al resolver el problema planteado destacamos los siguientes: - No identificar el uso del teorema central del límite. - Confundir la media poblacional y la media muestral. - Tomar la desviación típica con la varianza. - Tipificar de forma errónea. - No saber utilizar la tabla de la Normal. - No identificar la media muestral como mejor estimador de la media poblacional. - Expresión del intervalo de confianza. - No conocer la expresión de la media y la varianza muestral

Tabla 4. *Resúmenes estadísticos del número de objetos identificados por participante*

	Mínimo	Máximo	Media	D. típica	N. errores previsto
Errores interpretación	0	0	0	0	5
Errores conceptuales	1	5	1,9	1,1	4
Errores procedimentales	0	4	2,3	0,9	6
Número total de errores	1	8	4,2	1,5	15
Incorrectos	0	4	0,5	0,9	0

Nota: Fuente propia de la investigación.

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

Como en el caso de E12, algunos errores descritos como posibles (expresión del intervalo de confianza) no se aplican al problema analizado. Pero el número de errores incorrectamente citados fue pequeño, no llegando a uno de promedio por participante.

Errores de tipo conceptual

Cada participante cita, aproximadamente, dos errores de tipo conceptual, resultado mejor que el obtenido en López-Martín et al. (2019), donde el número medio de errores conceptuales mencionados no llega a 1 en cada problema. Por otro lado, se citó alguna vez, aunque con baja frecuencia, todos los errores previstos en el análisis a priori de la actividad. Los errores conceptuales más frecuentemente identificados fueron los siguientes:

- *Confundir la distribución de la media muestral y la distribución poblacional, o no identificar sus parámetros* (23 participantes). Se hace alusión a que el estudiantado podría no llegar a diferenciar la distribución correspondiente a la altura de los estudiantes y la distribución de la media muestral de una muestra de 25 estudiantes, cuando se recogen diferentes muestras sucesivas, es decir, al error apuntado por Harradine et al. (2011), debido a la dificultad de manejar en el mismo problema las dos distribuciones diferentes. En otros casos, se alude al error consistente en no calcular correctamente los parámetros de la distribución de la media muestral. Por ejemplo, en la respuesta de E13 presentada a continuación, el estudiante señala el olvido de dividir por el tamaño de la muestra en el cálculo de la desviación típica muestral. Este tipo de error fue identificado por Vallecillos (1999).

E13: Un posible fallo de los estudiantes sería el olvidar dividir por $\sqrt{n}$ a la hora de determinar los parámetros de la distribución que sigue la variable estudiada.

- *Confundir o aplicar erróneamente las propiedades de la distribución normal.* Por ejemplo, es necesario recordar la simetría de la distribución respecto a su media, e igualmente la simetría de la distribución tipificada N (0,1) respecto al origen de coordenadas, para posteriormente utilizar la tabla y hallar la probabilidad pedida. En el proceso de cálculo de la probabilidad una propiedad esencial es que el área bajo la curva normal (en realidad bajo cualquier función de densidad) es igual a la unidad. Esta propiedad fue citada tan solo por 13 futuros profesores o profesoras, a pesar de su importancia en la resolución del problema.

E14: Más concretamente, olvidar las propiedades o cometer errores asociados a una comprensión incompleta de la distribución normal.

- *Confundir estadístico y parámetro*. Identificamos que 11 estudiantes, como el presentado como ejemplo, señalan que se podrían confundir el estadístico media muestral y el parámetro media poblacional, error resaltado por Harradine et al. (2011) y que llevaría a suponer que la probabilidad pedida en el segundo apartado se refiere a la altura en la población y no a la altura media en diferentes muestras.

E15: Confundir la media poblacional y la media muestral.

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

- *Error al aplicar el teorema central del límite* o, como en el siguiente ejemplo, no reconocer que la distribución de la media muestral es también normal. Ello es en realidad debido, no al teorema central del límite, sino a que la población de partida es normal (6 casos).

E16: No conocer que la distribución de la media muestral es también igual que la distribución de la población que se estudia.

En consecuencia, los principales errores conceptuales identificados en nuestro análisis a priori de la tarea fueron reconocidos por algunos futuros profesores o profesoras de la muestra, pero en muy escaso número, teniendo en cuenta la cantidad de participantes en el estudio y en concordancia con los resultados de López-Martín et al. (2019), donde también el número de errores conceptuales citados por los participantes fue pequeño. Pero a diferencia de aquel estudio, en que algunos de los principales errores conceptuales previsibles no fueron identificados, en el nuestro todos han sido citados, aunque con muy poca frecuencia.

Errores de tipo procedimental

Finalmente, se podrían realizar errores de tipo aritméticos al completar los cálculos. En relación con los errores de tipo procedimental, que se producen al resolver la segunda parte del problema en la que se pide calcular una probabilidad, los más citados fueron los siguientes:

- *Errores de tipificación.* Se alude a las posibles dificultades de los estudiantes cuando lleven a cabo la tipificación de la variable media muestral, de forma que, mediante dicha operación se llegue a la distribución normal estándar N (0,1). Se indica, bien resumidamente o bien explicando en qué consiste la operación de tipificación, como ocurre con E17. En total son 30 los casos en que se citan errores de tipificación.

E17: Normalizar la distribución que siga la variable a tratar. Por ejemplo, si la variable sigue una normal de media 165 y desviación típica 10, su normalización se obtendría restándole la media a la variable y dividiendo el resultado por la desviación típica.

- *Errores por incorrecta lectura de la tabla de la distribución normal*, debido al hecho de que, una vez tipificada la variable, la probabilidad a calcular se transforma en P $(Z > -2{,}5)$, siendo Z una variable normal N (0,1). Pero la tabla de la distribución normal utilizada por los estudiantes proporciona los valores de su función de distribución P $(Z \leq k)$. Además, como la tabla proporcionada incluye únicamente los valores positivos de la variable, habría que aplicar la propiedad de simetría de la distribución normal, obteniendo P $(Z > -2{,}5)$ = P $(Z < 2{,}5)$, que se puede mirar directamente en la tabla proporcionada. Un total de 28 estudiantes indicaron este tipo de errores, con mayor o menor precisión en el lenguaje.

E18: Destacaría, en primer lugar, que se cometan errores en la búsqueda de las tablas de la normal, sobre todo, el no transformar la k (o hacerlo mal) para que sea positiva y la expresión a buscar en la tabla sea del tipo P $(Z \leq k)$.

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

- *Errores en el cálculo de probabilidades*. Como hemos indicado, para el cálculo de la probabilidad pedida es necesario aplicar algunas propiedades de la distribución normal, en concreto la simetría y la probabilidad del suceso contrario. Aplicar estas propiedades no es trivial, como se muestra en la investigación de Batanero et al. (2004) y en la de Onyancha y Ogbonnaya (2022). En nuestro estudio, esta dificultad fue solo citada por 15 participantes.

E19: pueden cometer errores en el cálculo de la probabilidad, si no saben cómo interpretar lo que se les pide o qué procedimiento aplicar.

- *Errores al operar las desigualdades*. Ocurre al despejar la variable por no tener un dominio correcto del trabajo con desigualdades. Este error citado por González y Ojeda (2017) fue descrito por 12 participantes de nuestra muestra. Además, 11 futuros profesores indicaron la posibilidad de cometer errores aritméticos en las operaciones realizadas. En el siguiente caso, se alude conjuntamente a los dos errores.

E20: También pueden cometer errores al realizar operaciones algebraicas y de despeje y de interpretación de signos negativos.

El error en la interpretación de los resultados obtenidos únicamente es citado por tres participantes. Finalmente, 12 participantes aluden a errores en la interpretación de las preguntas planteadas en el enunciado e identificación de los datos requeridos o la probabilidad requerida.

E21: Los estudiantes pueden errar al interpretar los datos dados en el enunciado (Ley Normal con una media y desviación típica dada).

Es decir, de nuevo se citan alguna vez, aunque con baja frecuencia todos los errores previstos en nuestro análisis. En resumen, los participantes en el estudio mostraron algún conocimiento de la faceta cognitiva relacionada con la distribución de la media muestral, aunque obviamente mejorable, por el escaso número de errores previstos en la tarea, a pesar de ser un tema de alta complejidad.

Conclusiones

La actividad descrita permitió evaluar algunos componentes del conocimiento didáctico-matemático (Godino, 2009; Godino et al., 2017; Pino-Fan y Godino, 2015; Pino-Fan et al., 2015) de profesores de matemáticas españoles en formación. Los resultados de la evaluación fueron muy buenos en lo que se refiere al conocimiento matemático común del tema, con la mayoría de los participantes resolviendo correctamente el problema propuesto. Las respuestas obtenidas superaron a las de Inzunsa y Jiménez (2013) con estudiantes del grado de matemáticas sobre el contraste de hipótesis, puesto que en su caso los participantes muestran un conocimiento fragmentado de los diversos conceptos que se involucran en las pruebas de hipótesis, mientras que en el presente estudio, al resolver el problema, los participantes relacionaron los diferentes conceptos, propiedades y procedimientos utilizados.

No obstante, se observaron algunos errores como, por ejemplo, la confusión de la distribución de la variable analizada en la población con la distribución muestral del estadístico, confusión que se ha descrito

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

entre los estudiantes (Begué et al., 2019; Espinel et al., 2007; Harradine et al. 2011; Vallecillos, 1999). Por tanto, este resultado es relevante para la formación de profesores y profesoras, pues, de lo contrario, algunos de ellos en el futuro podrían contribuir a transmitir este error a sus estudiantes.

Los participantes mostraron una competencia razonable de análisis de conceptos y procedimientos requeridos para solucionar la tarea propuesta, puesto que todos los previstos en el análisis a priori fueron citados por algún participante. Sin embargo, tuvieron mayor dificultad en reconocer las propiedades empleadas en la solución y el número medio de objetos (conceptos, propiedades y procedimientos) citados por cada participante fue bastante menor que la cantidad prevista.

También fue menor la competencia de análisis de los posibles errores que podrían cometer los estudiantes en la resolución de la tarea. Este resultado es comparable con el obtenido en la investigación de López-Martín et al. (2019), en la que un grupo de futuros profesores de educación secundaria y bachillerato españoles analizaron los posibles errores en dos problemas, uno de contraste de hipótesis y otro de intervalo de confianza. Por tanto, hay coincidencia en que los futuros profesores españoles que participan en las dos investigaciones no mostraron consciencia de la dificultad del tema para sus estudiantes ni previeron sus posibles errores.

En consecuencia, el estudio revela puntos de mejora en la formación de los futuros profesores españoles sobre la distribución de la media muestral, un contenido relevante para la posterior comprensión de otros contenidos propios de la inferencia, como el contraste de hipótesis o el intervalo de confianza.

Respecto al conocimiento común del contenido (Godino et al., 2017; Pino-Fan y Godino, 2015), dicha formación debiera enfatizar la diferencia entre las tres distribuciones de probabilidad que aparecen en el muestreo (Harradine et al., 2011; Vallecillos, 1999) y la diferencia entre estadístico y parámetro (Espinel et al., 2007; Harradine et al., 2011), pues el futuro profesorado no reconoce la posibilidad de errores de este tipo en su alumnado. Para ello, podría utilizarse la tecnología para analizar simulaciones basadas en la tarea presentada o en otras posibles, por ejemplo, las tareas de muestreo analizadas por Batanero et al. (2020), en las cuales se propone generar posibles resultados de muestras obtenidas con una población y el análisis de las distribuciones muestrales resultantes.

Es necesario, igualmente, realizar actividades formativas relacionadas con las facetas epistémica y cognitiva del conocimiento de la inferencia estadística, en las que se proponga el análisis de los objetos matemáticos y de los posibles errores de los estudiantes en tareas similares a la analizada en este trabajo. Finalmente, sería necesario investigar el conocimiento de los futuros profesores sobre el resto de facetas descritas en el modelo CDM (Godino, 2009; Godino et al., 2017; Pino-Fan y Godino, 2015; Pino-Fan et al., 2015), que no se han abordado en la presente investigación.

Financiamiento

Proyecto PID2019-105601GB-I00 / AEI / 10.13039/501100011033 y Grupo FQM126 (Junta de Andalucía).

Conflicto de intereses

Los autores declaran no tener algún conflicto de interés.

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

Declaración de la contribución de los autores

Todos los autores afirmamos que se leyó y aprobó la versión final de este artículo. El porcentaje total de contribución para la conceptualización, preparación y corrección de este artículo fue el siguiente: SVR: 30 %, CB: 30 %, NB: 20%, JGG: 20%.

Declaración de disponibilidad de los datos

Los datos que respaldan los resultados de este estudio serán puestos a disposición por el autor correspondiente [NB], previa solicitud razonable.

Referencias

Abu-Ghalyoun, O. (2021). Pre-service teachers' difficulties in reasoning about sampling variability. *Educational Studies in Mathematics*, *108*(3), 553-577. https://doi.org/10.1007/s10649-021-10067-8

Batanero, C., Begué, N., Borovcnik, M., & Gea, M. M. (2020). Ways in which high-school students understand the sampling distribution for proportions. *Statistics Education Research Journal*, *19*(3), 32-52.

Batanero, C., López-Martín, M. M., Gea, M. M., & Arteaga, P. (2018). Conocimiento del contraste de hipótesis por futuros profesores de educación secundaria y bachillerato. *Publicaciones*, *48*(2), 73-95. https://revistaseug.ugr.es/index.php/publicaciones/article/view/8334/7735

Batanero, C., Tauber, L. M., & Sánchez, V. (2004). Students' reasoning about the normal distribution. En D. Ben-Zvi, & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 257-276). Springer. https://doi.org/10.1007/1-4020-2278-6_11

Begué, N., Batanero, C., Ruiz, K., & Gea, M. M. (2019). Understanding sampling: A summary of the research. *Boletín de Estadística e Investigación Operative, BEIO, 35*(1), 49-78.

Beswick, K., & Goos, M. (2018). Mathematics teacher educator knowledge: What do we know and where to from here? *Journal of Mathematics Teacher Education*, *21*(5), 417-427. https://doi.org/10.1007/s10857-018-9416-4

Espinel, M. C., Ramos, R. M., & Ramos, C. E. (2007). Algunas alternativas para la mejora de la enseñanza de la inferencia estadística en Secundaria. *Números. Revista de Didáctica de las Matemáticas*, *67*, 15-23. https://ensciencias.uab.cat/article/view/v11-n3-mumbru/2418

Gal, I., & Geiger, V. (2022). Welcome to the era of vague news: a study of the demands of statistical and mathematical products in the COVID-19 pandemic media. *Educational Studies in Mathematics*, 111, 5-28. https://doi.org/10.1007/s10649-022-10151-7

Godino, J. D. (2009). Categorías de análisis de conocimientos del profesor de matemáticas. *Unión*, *20*, 13-31.

Godino, J., Batanero, C., & Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM - Mathematics Education, 39*(12), 127-135. https://doi.org/10.1007/s11858-006-0004-1

Godino, J., Batanero, C., & Font, V. (2019). The ontosemiotic approach: Implications for the prescriptive character of didactics. *For the Learning of Mathematics, 39*(1), 38-43.

Godino, J. D., Giacomone, B., Batanero, C., & Font, V. (2017). Enfoque ontosemiótico de los conocimientos y competencias del profesor de matemáticas. *Bolema*, *31*(57), 90-113. https://doi.org/10.1590/1980-4415v31n57a05

González, Y., & Ojeda, A. M. (2017). Comprensión de la distribución normal en bachillerato. En L. A. Serna (Ed.), *Acta Latinoamericana de Matemática Educativa* (pp. 207-217). Comité Latinoamericano de Matemática Educativa.

Haller, H., & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research*, *7*(1), 1-20.

Harradine, A., Batanero, C., & Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. En C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher education* (pp. 235-246). Springer. https://doi.org/10.1007/978-94-007-1131-0

Inzunsa, S., & Jiménez, J. V. (2013). Caracterización del razonamiento estadístico de estudiantes

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470

universitarios acerca de las pruebas de hipótesis. *Revista Latinoamericana de Investigación en Matemática Educativa*, *16*(2), 179-211. https://doi.org/10.12802/relime.13.1622

Kula, F., & Koçer, R. G. (2020). Why is it difficult to understand statistical inference? Reflections on the opposing directions of construction and application of inference framework. *Teaching Mathematics and its Applications, 39*(4), 248-265. https://doi.org/10.1093/teamat/hrz014

Llinares, S. (2018). Mathematics teacher's knowledge, knowledge-based reasoning, and contexts. *Journal of Mathematics Teacher Education*, *21*(1), 1-3. https://doi.org/10.1007/s10857-018-9399-1

Liu, Y., & Thompson, P. W. (2009). Mathematics teachers' understandings of proto-hypothesis testing. *Pedagogies: An International Journal*, *4*(2), 126-138. https://doi.org/10.1080/15544800902741564

López-Martín, M. M., Batanero, C., & Gea, M. M. (2019). ¿Conocen los futuros profesores los errores de sus estudiantes en la inferencia estadística*? Bolema: Boletim de Educação Matemática*, *33*, 672-693. https://doi.org/10.1590/1980-4415v33n64a11

Lugo-Armenta, J. G., & Pino-Fan, L.R. (2021). Inferential statistical reasoning of math teachers: experiences in virtual contexts generated by the COVID-19 pandemic**.** *Education Sciences, 11*, 363. https://doi.org/10.3390/educsci11070363

Makar K., & Rubin A. (2018) Learning about statistical inference. En D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International Handbook of Research in Statistics Education. Springer International Handbooks of Education*. https://doi.org/10.1007/978-3-319-66195-7_8

Mendenhall, W., Scheaffer, R. L., & Lyman, R. (2006). *Elementos de muestreo*. Paraninfo.

Ministerio de Educación, Cultura y Deporte. (2015). *Real decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la educación secundaria obligatoria y del bachillerato.* MECD.

Ministerio de Educación y Formación Profesional. (2022). *Real Decreto 217/2022, de 29 de marzo, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Secundaria Obligatoria.* MEFP.

Muñiz-Rodríguez, L., Alonso, P., Rodríguez-Muñiz, L. J., & Valcke, M. (2016). Is there a gap in initial secondary mathematics teacher education in Spain compared to other countries? *Revista de Educación*, *372,* 106–132. https://doi.org/10.4438/1988-592X-RE-2015-372-317

Onyancha, B. N., & Ogbonnaya, U. I. (2022). Undergraduate students' proficiencies in solving bivariate normal distribution problems. *Journal of Pedagogical Sociology and Psychology, 4*(1), 20-34. https://doi.org/10.33902/JPSP.202212259

Pino-Fan, L. R., & Godino, J. D. (2015). Perspectiva ampliada del conocimiento didáctico-matemático del profesor. *Paradigma*, *36*(1), 87-109.

Pino-Fan, L., Assis, A., & Castro, W. F. (2015). Towards a methodology for the characterization of teachers' didactic-mathematical knowledge. *EURASIA Journal of Mathematics, Science and Technology Education*, *11*(6), 1429-1456. https://doi.org/10.12973/eurasia.2015.1403a

Scheiner, T., Montes, M. A., Godino, J. D., Carrillo, J., & Pino-Fan, L. R. (2019). What makes mathematics teacher knowledge specialized? Offering alternative views. *International Journal of Science and Mathematics Education*, *17*(1), 153-172. https://doi.org/10.1007/s10763-017-9859-6

De Vetten, A., Schoonenboom, J., Keijzer, R., & Van Oers, B. (2018). Pre-service primary school teachers' knowledge of informal statistical inference. *Journal of Mathematics Teacher Education*, *22*(6), 639–661. https://doi.org/10.1007/s10857-018-9403-9

Vallecillos, A. (1999). Some empirical evidence on learning difficulties about testing hypotheses. *Bulletin of the International Statistical Institute: Proceedings of the Fifty second Session of the International Statistical Institute* (pp. 201-204). International Statistical Institute https://www.stat.auckland.ac.nz/~iase/publications/5/vall0682.pdf

Watson, J. M., & Callingham, R. (2013). Likelihood and sample size: The understandings of students and their teachers. *The Journal of Mathematical Behavior*, *32*(3), 660–672. https://doi.org/10.1016/j.jmathb.2013.08.003

Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics, 96*(1), 33–47. https://doi.org/10.1007/s10649-017-9764-5

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr

DOI: http://dx.doi.org/10.15359/ru.37-1.3
E-ISSN: 2215-3470
CC: BY-NC-ND

UNICIENCIA Vol. 37, N°. 1, pp. 1-20. January-December, 2023 • www.revistas.una.ac.cr/uniciencia • revistauniciencia@una.cr