http://doi.org/10.14393/ER-v31e2024-03

Comprensión que muestran futuros profesores de Matemática y estadística de los intervalos de confianza sobre la media

Nicolás Sánchez Acevedo[1]

RESUMEN

Los intervalos de confianza son un aspecto importante en la inferencia estadística. Actualmente este tema se ha incluido en el curriculum de la formación de profesores de Matemática y los programas de estudio de secundaria en Chile. De acuerdo con esto, se explora en la comprensión que manifiestan futuros profesores de Matemática y Estadística sobre los intervalos de confianza sobre la media. Se ha adoptado una metodología cualitativa y de tipo exploratoria-descriptiva. Para ello se analizan las respuestas de 11 futuros profesores de Matemática y Estadística, que llevaban el curso de inferencia estadística, a partir de un cuestionario de respuesta abierta con dos actividades sobre intervalos de confianza para la media y diferencia de medias. Los resultados evidencian confusión en el cálculo de intervalos de confianza para la media, interpretaciones deterministas y errores en la determinación de cuantiles Z o T. A pesar de ello, la gran mayoría de los futuros profesores es consciente de los supuestos que subyacen a la construcción de estos intervalos de confianza. Estos resultados son relevantes, tanto para los futuros profesores, como formadores de profesores de Matemáticas, dado los pocos estudios sobre este tema en el contexto del profesorado.

PALABRAS CLAVE: Intervalos de confianza; Futuros profesores de matemática; Estadísticas; Cualitativo.

[1] Dr(c) en Ciencias, área de Matemática Educativa (CICATA-IPN, México). Universidad Alberto Hurtado/Universidad Central de Chile, Santiago, Chile. Orcid: https://orcid.org/0000-0002-0665-6102. *E-mail*: nsanchez@uahurtado.cl.

http://doi.org/10.14393/ER-v31e2024-03

Compreensão de futuros professores de Matemática e Estatística em relação aos intervalos de confiança para a média

RESUMO

Os intervalos de confiança, assim como os testes de hipóteses, desempenham um papel importante na inferência estatística. Atualmente, este tema foi incorporado no currículo da formação de professores de Matemática e nos programas de estudo do ensino secundário no Chile. Nesse sentido, investigamos a compreensão demonstrada pelos futuros professores de Matemática e Estatística em relação aos intervalos de confiança para a média. Adotamos uma metodologia qualitativa de natureza exploratória-descritiva. Para isso, analisamos as respostas de 11 futuros professores de Matemática e Estatística, que estavam cursando um curso de inferência estatística, por meio de um questionário de resposta aberta com duas atividades relacionadas aos intervalos de confiança para a média e diferenças de médias. Os resultados revelam confusão no cálculo dos intervalos de confiança para a média, interpretações determinísticas e erros na determinação dos quantis Z ou T. No entanto, a grande maioria dos futuros professores está ciente das premissas subjacentes à construção desses intervalos de confiança. Esses resultados são relevantes tanto para os futuros professores quanto para os formadores de professores de Matemática, dado o escasso número de estudos sobre esse tema no contexto do ensino.

PALAVRAS-CHAVE: Intervalos de confiança; Futuros professores de matemática; Estatísticas; Qualitativo.

Future Mathematics and Statistics teachers' understanding of confidence intervals about the arithmetic average

ABSTRACT

Confidence intervals, like hypothesis tests, play a crucial role in statistical inference. Currently, this topic has been incorporated into the curriculum for mathematics teacher training and secondary school programs in Chile. In line with this, we investigate the understanding expressed by future Mathematics and Statistics teachers regarding

http://doi.org/10.14393/ER-v31e2024-03

confidence intervals for the mean. We have adopted a qualitative, exploratory-descriptive methodology for this purpose. To achieve this, we analyze the responses of 11 future Mathematics and Statistics teachers who were taking a course in statistical inference, using an open-response questionnaire with two activities related to confidence intervals for the mean and differences in means. The results reveal confusion in the calculation of confidence intervals for the mean, deterministic interpretations, and errors in determining Z or T quantiles. Nevertheless, most future teachers are aware of the assumptions underlying the construction of these confidence intervals. These findings are relevant both for future teachers and for mathematics teacher educators, given the limited research on this topic in the context of teacher education.

KEYWORDS: Confidence intervals; Future Mathematics teachers; Statistics; Qualitative.

* * *

Introducción

El aprendizaje y la enseñanza de la estadística ha tomado un papel relevante por su rol en la construcción de ciudadanos críticos (MINEDUC, 2015; NCTM, 2000). La importancia que ha adquirido la estadística en la sociedad, debido a la masiva cantidad de información disponible, se ha traducido en que diversos investigadores se han interesado por desarrollar investigación en diversos campos de aplicación profesional, como las ciencias básicas, ingeniería, bachillerato, profesores en formación y en ejercicio, con el fin de profundizar en las dificultades y la comprensión sobre diferentes conceptos y procedimientos estadísticos. La intención de estos trabajos es obtener evidencia de estas dificultades, y con ello, proponer vías alternativas en la construcción de situaciones de enseñanza y aprendizaje para promover una mayor comprensión conceptual en esta área del conocimiento (SHAUGHNESSY, 2007), y desarrollar el pensamiento estadístico.

Como parte del pensamiento estadístico, la inferencia cumple un papel relevante en el comportamiento de características poblacionales. En este sentido, los intervalos de confianza son un mecanismo elemental para realizar estimaciones de las características de la población bajo estudio y, son un complemento a los contrastes de hipótesis (Roldán, Batanero; Álvarez-Arroyo, 2020). Así mismo, tanto los intervalos de confianza, como los contrastes de hipótesis, son una parte intrínseca en la forma de pensar la estadística como una disciplina metodológica.

Hoy en día, el tratamiento de los intervalos de confianza se promueve a nivel profesional para informar y comprender diversos fenómenos (Coulson, Healey, Fidel; Cumming, 2010; Yaremko, Harari, Harrison; Lynn, 2013), mucho más que el de los contrastes de hipótesis.

Dada la importancia que han cobrado los intervalos de confianza [y los contrastes de hipótesis], como parte del desarrollo de la alfabetización estadística (GAL, 2002), es que en la actualidad diversos curriculum han incluido estos contenidos en las directrices de enseñanza, tanto a nivel escolar como universitario (e.g. Franklin et al., 2007; MEC, 2015: MINEDUC, 2021a; MINEDUC, 2021b NCTM, 2000).

En Chile, la incorporación de la unidad de Estadística y Probabilidad (desde el año 2009), como parte del curriculum de matemática (MINEDUC, 2012; MINEDUC, 2021a; MINEDUC, 2021b), incluye contenidos de estadística descriptiva (primeros años escolares), hasta cubrir los temas relacionados con el concepto de distribución muestral, estimación (estimación puntual e intervalar) y contraste de hipótesis, para los casos de la media poblacional, asumiendo varianza conocida (MINEDUC, 2021b). La inclusión de estos temas en el curriculum, para el aprendizaje de los estudiantes, generó la necesidad de reformar los currículos en la formación de los futuros profesores de Matemática (en Chile), para que la formación recibida se adecue a las necesidades de la sociedad. Es por esto, que la mayoría de las carreras universitarias formadoras de profesores, cuentan con al menos dos cursos de estadística y, algunas otras, que profundizan en esta área, incluyen hasta cuatro cursos disciplinares[2], dentro de los que cuentan los de didáctica de la estadística y probabilidad, para fortalecer los aspectos de la estadística y su enseñanza.

La importancia de promover aprendizajes sobre la estadística, en general, y de los intervalos de confianza, en particular, ha sido reconocida en el curriculum chileno de educación secundaria (MINEDUC, 2021a) en el programa de estudio de 3° (16 a 17 años) y 4° medio (17 a 18 años). En el que se declara: (i) “que los estudiantes comprendan el uso de la estadística inferencial y cómo estimar parámetros de una población a partir de estadísticos de muestras de esa población” (MINEDUC, 2021a, p. 159) y, (ii) “que los estudiantes puedan obtener intervalos de confianza para realizar inferencias sobre la media de una población en diferentes contextos [...] según el nivel de confianza requerido” (MINEDUC, 2021a, p. 165).

Para promover el aprendizaje de los estudiantes de acuerdo con los propósitos de aprendizaje declarados, se proponen los *Estándares Pedagógicos y Disciplinarios para Carreras de Pedagogía en Matemáticas*[3] (MINEDUC, 2021b), declarando que el futuro profesor de Matemática debe ser capaz de “Implementar discusiones en clase para monitorear los diversos niveles de razonamiento y las dificultades que presentan sus estudiantes al interpretar los intervalos de confianza en problemas de inferencia estadística” (MINEDUC, 2021b, p. 89)

Tanto el programa de estudio de 3° y 4° medio (MINEDUC, 2021a), como los Estándares Pedagógicos y Disciplinarios para Carreras de Pedagogía en Matemática (MINEDUC, 2021b) muestran una coherencia entre lo que debe saber el profesor de Matemática y lo que debe promover en los estudiantes para la comprensión de estos contenidos. Se destaca de los

[2] En Chile, dentro de las carreras de formación de profesores de Matemática actualmente vigentes en las Instituciones de educación superior, sólo una de ellas que cuenta con una profundización disciplinar en la línea de estadística. Esta carrera de formación del profesorado, además de incluir los cursos de Estadística descriptiva, probabilidades e inferencia estadística, incluye un curso llamado modelos estadísticos, donde se abordan modelos lineales clásicos (RLS y RLM) y algunos modelos no lineales para contexto de enseñanza y aprendizaje.

[3] En Chile, la formación de profesores, en las diferentes disciplinas, esta ordenada por los estándares de formación inicial docente (https://estandaresdocentes.mineduc.cl/), los cuales “tiene por como propósito el fortalecer la formación de docentes y asegurar profesores y profesoras mejor preparados/as para enfrentar los grandes desafíos que representa esta sociedad del conocimiento, dinámica y globalizada” (MINEDUC, 2021, p. 3).

Estándares Pedagógicos y Disciplinarios para Carreras de Pedagogía en Matemáticas[4] (MINEDUC, 2021b), la ausencia en la explicitación de los supuestos que se deben cumplir en la construcción de los intervalos de confianza para los parámetros poblacionales μ y p.

Dado que los intervalos de confianza son un tema relativamente nuevo en el curriculum chileno, esto demanda poder explorar en la comprensión de futuros profesores de Matemática para contextos de enseñanza. La investigación existente en este ámbito es escasa, tanto en la formación del profesorado, como en profesores en ejercicio, pues gran parte de los estudios que se han realizado sobre intervalos de confianza se han centrado en estudiantes de ingeniería o profesionales de otras áreas (como se describe en el apartado de la fundamentación teórica).

En este sentido, el objetivo de esta investigación de tipo exploratoria es evaluar la comprensión que manifiestan futuros profesores de Matemática y Estadística sobre los intervalos de confianza.

Fundamentos teóricos

En la actualidad, las investigaciones en educación estadística han aumentado significativamente, pero en particular, los estudios relacionados con la comprensión sobre los intervalos de confianza son escasas. Aquellas realizadas, se han centrado en investigadores profesionales que usan IC, y otras, en estudiantes universitarios y profesionales.

Uno de los trabajos que incluye profesionales vinculados con la enseñanza es el de Behar (2001), quien exploró la comprensión de los intervalos en 47 profesionales estadísticos-profesores y 297 estudiantes de ingeniería. Se les aplicó un cuestionario sobre IC y test de hipótesis. De las respuestas, un 29% de los profesionales estadísticos y un 59% de los

[4] En Chile, la formación de profesores, en las diferentes disciplinas, esta ordenada por los estándares de formación inicial docente (https://estandaresdocentes.mineduc.cl/), los cuales "tiene por como propósito el fortalecer la formación de docentes y asegurar profesores y profesoras mejor preparados/as para enfrentar los grandes desafíos que representa esta sociedad del conocimiento, dinámica y globalizada" (MINEDUC, 2021, p. 3).

http://doi.org/10.14393/ER-v31e2024-03

estudiantes consideran el nivel de confianza como un porcentaje en el que caen los datos poblacionales y no el parámetro. Algunos errores estuvieron en dar interpretaciones bayesianas, establecer relaciones de proporcionalidad directa entre el ancho del intervalo y tamaño de la muestra, como también, no ven la utilidad de los intervalos de confianza para la toma de decisiones sobre las hipótesis, al no visualizar un IC como un conjunto de valores posibles del verdadero parámetro poblacional.

En el contexto de los trabajos con investigadores profesionales, está el trabajo de Cumming, William y Fidler (2004). Esta investigación se llevó a cabo con 134 investigadores que desarrollaban trabajos con el uso de intervalos de confianza. En este contexto se les entregó un intervalo con un 95% de confianza para estimar la media poblacional, y se les pidió que dieran valores factibles para la media en nueve experimentos. De los 134 investigadores, el 78% manifestó que el intervalo de confianza contendría al verdadero parámetro en el 95% de las veces (probabilidad de replicación, ROLDÁN et al., 2020), lo que evidencia algunas concepciones erróneas al creer que los límites de un IC son constantes.

Debido al abuso y mala interpretación que se hace, en ocasiones, sobre las pruebas de hipótesis, Hoekstra, Morey, Rouder y Wagenmakers (2014) pidieron a 120 investigadores y 442 estudiantes del área de psicología que evaluaran la veracidad de seis afirmaciones que involucraban la interpretación de IC, las que en su totalidad eran falsas. No obstante lo anterior, tanto los investigadores, como los estudiantes, afirmaron (en media), tres o más de las aseveraciones planteadas, lo que evidencia una comprensión inadecuada sobre los I.C. Dentro de los aspectos delicados de este trabajo, es que pocos investigadores manifestaron una comprensión por sobre la de los estudiantes, considerando que estos últimos no habían tenidos cursos sobre inferencia estadística. Estos resultados muestran que muchos investigadores no conocen la interpretación correcta de un IC, lo que constituye un problema *grave* a la hora de tomar decisiones con base en los datos.

Rashidah, Razak, Baharun y Arul (2018) realizaron una investigación sobre intervalos de confianza, pero en la estimación de los parámetros de un modelo de regresión lineal simple. El estudio se llevó a cabo con 197 estudiantes de segundo año de una diplomatura en estadística y en ciencias actuariales, a los que se les aplicó un cuestionario y se analizaron las respuestas de los estudiantes sobre el cálculo de los IC para la pendiente y su posterior interpretación. De los 197 estudiantes, el 48% logró calcular correctamente el IC para el parámetro de la pendiente β_0 y el 68,5% de los estudiantes entregó interpretaciones erróneas sobre la pendiente β_1 del modelo estimado. El 68,5% de los estudiantes proporcionó interpretaciones incorrectas que demostraron su incapacidad para comprender el concepto de pendiente de regresión. Dentro de los errores detectados se encontró que los estudiantes no identificaron adecuadamente lo grados de libertad y no evaluaron correctamente el error estándar de la pendiente.

En al ámbito de las investigaciones sobre IC en estudiantes de diversos niveles, esta la de Yañez y Behar (2000), quienes realizaron un estudio en el que muestran posibles explicaciones a las interpretaciones erróneas que realizan estudiantes y profesores con respecto de los intervalos y niveles de confianza por medio de un estudio epistemológico y entrevistas clínicas. Los resultados encontrados mostraron que las interpretaciones erróneas se deben, por una parte, a los procesos en el tratamiento estadístico-algebraico de la construcción de los intervalos de confianza, y por otra, a la extrapolación de la probabilidad del nivel de confianza de que el verdadero parámetro este en el IC dado.

Batanero y Olivo (2007) evaluaron la comprensión que tienen estudiantes de ingeniería sobre los intervalos de confianza en una muestra de 48 estudiantes. La información fue recogida a partir de la construcción y validación de un cuestionario de opción múltiple. Los resultados mostraron una alta comprensión de los intervalos de confianza a nivel conceptual, pero se presentaron errores procedimentales y conceptuales emergentes como la confusión de la distribución muestral, no determinar el valor crítico, o confundir el estadístico con el parámetro.

Olivo, Batanero y Díaz (2008) llevaron a cabo una investigación para evaluar la comprensión de estudiantes universitarios de carrera de ingeniería sobre intervalos de confianza. Se aplicó un cuestionario que incluía 12 ítems y dos situaciones aplicadas abiertas a una muestra de 252 estudiantes con estudios sobre los IC. Los resultados mostraron una falta de comprensión en la aleatoriedad de los extremos de los intervalos, no comprenden que la anchura del intervalo se puede deber a diversos factores, no seleccionan la distribución muestral adecuada de acuerdo con la información proporcionada y, se presentan errores en la determinación de los valores críticos.

El trabajo de Henriques (2016) estudió las dificultades en la comprensión de los intervalos de confianza en una muestra de 33 estudiantes de segundo año de formación (en una escuela naval), a través de una prueba escrita como parte de un curso de introducción a la estadística. La prueba estuvo conformada por 4 preguntas de selección múltiple y cinco de respuesta abierta sobre cuestiones sobre I.C. Si bien los resultados mostraron que los estudiantes comprenden el concepto de intervalo de confianza, se evidenciaron algunas dificultades, tanto procedimentales, como conceptuales, por ejemplo, la comprensión de la definición de un IC, la interpretación descriptiva del intervalo de confianza, la selección de la distribución muestral (en particular par la varianza), como también, la determinación de los valores críticos.

Crooks, Bartel y Alibali (2019) llevaron a cabo un estudio con estudiantes de pre y posgrado en el área de sociología que evaluó la comprensión conceptual que tienen sobre los intervalos de confianza. Dentro de los resultados encontrados, se evidenciaron errores conceptuales en su tratamiento, una falta en la conexión entre los conceptos de IC y la estimación de la media muestral y, una falta de relación entre los IC hacia las pruebas de una hipótesis nula. A partir de estos resultados, se sugieren profundizar en el tratamiento de elementos básicos de IC.

http://doi.org/10.14393/ER-v31e2024-03

Roldán et al. (2020) realizaron una investigación que tuvo por objetivo evaluar la comprensión sobre los intervalos de confianza por estudiantes de bachillerato a través de un cuestionario que incluyó dos tipos de ítems, de selección múltiple y de respuesta abierta (adaptado de otras investigaciones). Dentro de los resultados obtenidos, se evidenció que hay una limitada comprensión de los intervalos de confianza para la estimación de un parámetro. Así mismo, no más de un 40% de los estudiantes evaluados (58) lograron desarrollar satisfactoriamente las actividades de manera completa.

De acuerdo con las investigaciones reportadas y algunas otras (e.g. Andrade; Fernández, 2016; Reaburn, 2014), es posible evidenciar que la investigación sobre la comprensión de los intervalos de confianza en el contexto de futuros profesores de matemática es escasa, por no decir nula, lo que abre un espacio para explorar en dicho contexto, que es lo que se propone esta investigación.

Un acercamiento al concepto de Intervalo de Confianza

Uno de los aspectos relevantes en estadística es convertir datos en información a partir de una muestra y hacer inferencias acerca de una o varias características de la población para tomar decisiones. Considerando que las poblaciones se caracterizan por medidas descriptivas llamadas parámetros, la investigación en estadística busca *estimar* el valor de estos parámetros (Wackerly; Mendenhall; Scheaffer, 2010).

En este sentido, los intervalos de confianza no son un concepto de fácil comprensión. En el caso de la interpretación frecuentista, se establece que el IC del $(100-\alpha)\%$ permite caracterizar un proceso de generación de intervalos, en la cual, la repetición del proceso, el $(100-\alpha)\%$ de los intervalos generados contienen el verdadero parámetro desconocido. [...], el tratamiento habitual de los intervalos de confianza, a menudo, comienza con una distribución muestral de un estadístico, como, por ejemplo, la media muestral (Engel, 2010).

En general, en el caso de la inferencia clásica [paramétrica], se asumen estimadores insesgados, por ejemplo, para μ, p, $\mu_1 - \mu_2$, $p_1 - p_2$ en los que se asume que, para muestras grandes, los estimadores puntuales tienen distribuciones muestrales aproximadamente normales con errores estándar particulares[5] y, además, se supone independencia de las muestras. Es decir, para muestras grandes se tiene $Z = \frac{\hat{\theta}-\theta}{\sigma_{\hat{\theta}}}$ tiene aproximadamente una distribución normal estándar, que funciona como una cantidad pivotal para el intervalo de confianza del parámetro objetivo θ.

Para el caso de esta investigación mostramos de manera amplia, considerado los supuestos anteriores, los intervalos de confianza para la media (con σ desconocida), y para la diferencia de medias en el caso de varianzas desconocidas y desiguales. En Wackerly et al. (2010, pp. 425-428) se describen ambas estimaciones intervalares de la siguiente manera:

- Para un intervalo de confianza de μ cuando se desconoce σ se supone una muestra aleatoria $X_1, X_2, \ldots, X_n$ de una población normal con media $\bar{X}$ y varianza S^2 muestrales estimadas. El intervalo de confianza para la media μ de la población, cuando $V(Y_i) = \sigma^2$ es desconocida y el tamaño de la muestra es pequeño llevan a usar la cantidad pivote $T = \frac{X-\mu}{S/\sqrt{n}}$, que tiene una distribución t con $(n-1)$ grados de libertad. Con valores adecuados para $t_{\alpha/2}$ y $-t_{\alpha/2}$ se puede plantear $P\left(-t_{\alpha/2} \leq T \leq t_{\alpha/2}\right) = 1 - \alpha$ y haciendo algunos arreglos algebraicos se obtiene el IC para μ con $\bar{X} \pm t_{\alpha/2,\ n-1} \cdot \left(\frac{S}{\sqrt{n}}\right)$. Cabe destacar que, una vía alternativa para la estimación intervalar de μ, se puede usar el teorema del límite central (TLC), el que plantea que si tenemos una sucesión $X_1, X_2, \ldots, X_n$ de variables aleatorias que provienen de una muestra de una distribución con media μ y varianza σ^2, para n suficientemente grande $(n \geq 30)$, se tiene que la distribución muestral de $\bar{X}$ es aproximadamente normal con media

[5] Ver Wackerly et al. (2010, p. 397)

$\mu_{\bar{X}} = \mu$ y $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ (Devore, 2012, p. 226). En este trabajo, la respuesta por TLC no es considera que no es parte de los objetivos del curso de inferencia para futuros profesores de Matemáticas.

- En Walpole et al. (2012, pp. 289-290) se describe el intervalo de confianza para la diferencia de medias $\mu_1 - \mu_2$ en el caso de tener varianzas desconocidas y desiguales de la siguiente manera: Si $\bar{X}_1$ y s_1^2 y, $\bar{X}_2$ y s_2^2 son las medias y varianzas de muestras aleatorias independientes de tamaños n_1 y n_2 respectivamente, tomadas de poblaciones aproximadamente normales con varianzas desconocidas y diferentes, un intervalo de confianza aproximado del $100(1-\alpha)\%$ para $\mu_1 - \mu_2$, está dado por $\bar{x}_1 - \bar{x}_2 \pm t_{\nu,\ \alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$. Los grados de libertad ν se aproximan[6].

Metodología

De acuerdo con el objetivo planteado en esta investigación, este trabajo se apoya en una metodología cualitativa (Stake, 2007) de tipo exploratoria-descriptiva (Pérez-Serrano, 1994; Hernández; Fernández; Baptista, 2010), puesto que nos interesa adentrarnos en la comprensión detallada que muestran un grupo de futuros profesores de Matemática y Estadística de una Universidad de Chile sobre intervalos de confianza sobre la media y diferencia de medias. Para analizar las respuestas de los futuros profesores nos apoyamos en el análisis de contenido (Noguero, 2002), siguiendo los siguientes pasos: (I) se seleccionan las actividades y las consignas respectivas a analizar. Esto debido a que cada respuesta se desglosa en partes relacionadas con los elementos de comprensión sobre IC; (II) a partir de los elementos sobre comprensión de IC discutidos en los

[6] $\nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left[\left(s_1^2/n_1\right)^2/(n_1-1)\right] + \left[\left(s_2^2/n_2\right)^2/(n_2-1)\right]}$ que son los grados de libertad estimados.

http://doi.org/10.14393/ER-v31e2024-03

fundamentos teóricos, se analizan las respuestas de los futuros profesores (logrado, parcialmente logrado, no logrado o no responde) y, (III) se realizan el análisis descriptivo de los datos realizando tablas para resumir y sistematizar la información obtenida.

Los futuros profesores y el contexto de formación

La población (N=11) de profesores fue intencional, y se accedió a ella por la facilidad de acceso a los estudiantes que realizaban el curso de Inferencia Estadística y la voluntad de colaborar en la investigación. Los futuros profesores, 3 mujeres (27,3%) y 8 hombres (72,7%), cursaban el 8° semestre de la carrera de Pedagogía en Matemática y Estadística en una institución de educación superior privada formadora de profesores de Chile. Los estudiantes, al momento de responder el cuestionario de respuesta abierta (como parte de una evaluación de proceso), habían cursado las asignaturas de Estadística Descriptiva y Probabilidades. El curso de Inferencia Estadística es prerrequisito del último curso de la línea de estadística llamado Modelos Estadísticos. Esta carrera (como ya se mencionó) es la única carrera de formación de profesorado de Matemática en Chile que profundiza en la línea de estadística y su didáctica.

El curso de Inferencia Estadística (donde recogimos los datos) se compone de tres grandes unidades de aprendizaje: (I) concepto de estimación y métodos de estimación puntual, donde se discuten los métodos de estimación de parámetros (máxima verosimilitud y momentos) y las propiedades de insesgamiento, eficiencia y consistencia, como también, el sesgo y error cuadrático medio; (II) estimación por intervalos de confianza, donde se muestra la construcción de intervalos de confianza para el caso clásico, donde se trabajan aspectos procedimentales y conceptuales de los intervalos de confianza para la media, la diferencia de medias, la proporción, diferencia de proporciones y, varianza; finalmente, (III) test de hipótesis, los aspectos conceptuales, tipos de errores (I y II) y los test de hipótesis clásicos.

Las actividades sobre intervalos de confianza

Como parte de una evaluación de proceso, se aplicó a los 11 estudiantes (población bajo estudio) un cuestionario con dos actividades de respuesta abierta sobre estimación por medio de intervalos de confianza para el caso de la media (μ) y la diferencia de medias ($\mu_1 - \mu_2$), en ambos casos para varianzas desconocidas.

Los ítems del cuestionario de respuesta abierta contenían consignas con el objeto de evaluar las estrategias de resolución, la argumentación de los estudiantes y su conocimiento sobre los supuestos que permiten la construcción de los intervalos. Todo ello, considerando los resultados de aprendizaje del curso y los estándares de formación inicial docente de Matemáticas (MINEDUC, 2021b).

Siguiendo el método de selección de actividades de Olivo et al. (2008), las dos actividades seleccionadas responden a los objetivos del curso de inferencia estadística. La actividad 1, tomada de Wackerly et al. (2010, p. 412) evalúa el conocimiento procedimental en la construcción y estimación del intervalo de confianza para la media de una población normal con varianza desconocida y, la comprensión conceptual de acuerdo con la interpretación y los supuestos a considerar. La actividad 2, tomada de Wackerly et al. (2010, p. 440) evalúa la comprensión sobre los supuestos que son necesarios para calcular intervalos de confianza para diferencias de medias y, evaluar el conocimiento procedimental y conceptual para el intervalo de confianza construido.

Actividad 1. Se registraron los tiempos de compra de n = 64 clientes seleccionados al azar en un supermercado local. El promedio y varianza de los 64 tiempos de compra fueron 33 minutos y 256 minutos, respectivamente. (a) Estime μ, el verdadero promedio de tiempo de compra por cliente, con un coeficiente de confianza de $1 - \alpha = 0{,}90$. (b) Interprete el intervalo de confianza en el contexto de la situación y, (c) ¿Qué debe suponer para la construcción del IC?

Respuestas esperadas

(a) Asumiendo un coeficiente de confianza de 90% se debe considerar que la varianza es desconocida, por lo tanto, lo que se tiene es una estimación de la varianza poblacional. Considerando un $n = 64$, se entrega un $\bar{x} = 33\ min.$ y una $s^2 = 256\ min.$ Usando $\bar{X} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$ bajo una distribución t-Student con $n - 1\ gl = 63\ gl$, la estimación intervalar para μ queda $33 \pm 1{,}67 \cdot \left(\frac{16}{8}\right)$, es decir, $29{,}66 < \mu < 36{,}34$. **(b)** Una interpretación adecuada seria: se espera que con un 95% el verdadero parámetro poblacional de la media se encuentre entre 29,66 min. y 36,34 min., lo que implica que, en el muestreo, por ejemplo, el 10% de las muestras de tamaño $n = 64$ no contendrán al verdadero tiempo promedio μ de la población. **(c)** Se espera que los futuros profesores (aspectos discutidos en clases de acuerdo con los supuestos a considerar) reflexionen que el uso de este intervalo de confianza hace necesaria una estimación de la varianza, por las características de la situación planteada. Además, deben considerar que las muestras de variables aleatorias X_i provienen de distribuciones normales, o en su defecto el tamaño es suficientemente grande y, que cada una de ellas es independiente e idénticamente distribuida, este último caso, podría asumir la construcción del intervalo aplicando el TLC.

Actividad 2. Dos métodos para enseñar a leer se aplicaron a dos grupos de niños de primaria seleccionados aleatoriamente y se compararon con base en un examen de comprensión de lectura al final del período de enseñanza. Las medias muestrales y varianzas calculadas a partir de las calificaciones de examen se muestran en la siguiente tabla.

Estadístico	**Método 1**	**Método 2**
Número de niños por grupo	11	14
$\bar{x}$	64	69
s^2	52	71

Se pide (a) ¿Qué suposiciones son necesarias? (b) Encuentre un intervalo de confianza de 95% para $(\mu_1 - \mu_2)$ e interprete el intervalo en el contexto de la situación presentada.

Respuestas esperadas

(a) En el caso de esta consigna, se espera que los estudiantes den cuenta de que la estimación del intervalo de confianza hace necesario saber que las muestras de las dos poblaciones son aleatorias e independiente. Además, que provienen de distribuciones con ley aproximadamente normal, donde no se puede asumir igualdad de varianzas. **(b)** Con base en las suposiciones de (a) se evalúa en la expresión $\bar{x}_1 - \bar{x}_2 \pm t_{v,\ \alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ para estimar la verdadera diferencia de medias. Se calcula la estimación de los grados de libertad para el caso de varianzas desiguales $v \approx 22{,}79 = 23$ y con ello tenemos $(64 - 69) \pm 2{,}069 \cdot \sqrt{\frac{52}{11} + \frac{71}{14}}$, y resolviendo, se obtiene $-11{,}476 < \mu_1 - \mu_2 < 1{,}477$. Con este intervalo, una posible interpretación sería: se espera que con una confianza del 95%, la verdadera diferencia de medias se encuentre entre -11,476 y 1,477, o también, que la verdadera diferencia de las calificaciones del examen de comprensión lectora se encuentra entre dichos valores. Posiblemente, algunos estudiantes podrían notar la relación con las pruebas de hipótesis al mencionar que los rendimientos de un grupo parecerían ser mejor que las del otro grupo.

Resultados

Mostramos en la Tabla 1 el porcentaje de logro a partir de las respuestas de los futuros profesores en relación con las consignas 1a, 1b, 1c de la actividad 1. El nivel de logro de los estudiantes se clasificó en: logrado, parcialmente logrado, no logrado y no responde de acuerdo con: la estimación del parámetro μ, la interpretación del intervalo de confianza y, las suposiciones para la construcción.

TABLA 1: Porcentajes de logro de la actividad 1.

Consigna por actividad	**Logrado**	**Parcialmente logrado**	**No logrado**	**No responde**
1a	*0%*	81,82%	18,18%	0%
1b	*0%*	81,82%	18,18%	0%
1c	72,73%	18,18%	9,09%	0%

Fuente: elaboración propia.

Se puede apreciar de la Tabla 1 que todos los profesores responden las consignas planteadas (1a, 1b y 1c). Además, se puede ver que el 72,73% de los futuros profesores conoce los supuestos que son necesarios para la construcción de IC para la media μ, que es un aspecto relevante en el contexto de futuros de profesores de Matemáticas. Así mismo, un 18,18% no logra construir el IC y no realiza la interpretación adecuada del IC.

En la Tabla 2, se muestra el porcentaje de logro de las consignas 2a y 2b de la actividad 2, de acuerdo con los niveles de respuesta antes especificados (logrado, parcialmente logrado, no logrado y no responde) para las suposiciones necesarias para la construcción del IC de la diferencia $(\mu_1 - \mu_2)$ y el procedimiento (proc.) e interpretación (int.) del IC.

TABLA 2: Porcentajes de logro de la actividad 2.

Consigna por actividad	**Logrado**	**Parcialmente logrado**	**No logrado**	**No responde**
2a	72,73%	0%	27,27%	0%
2b(proc.)	*45,45%*	27,27%	18,18%	9,1%
2b(int.)	*18,18%*	*45,45%*	18,18%	18,18%

Fuente: elaboración propia.

En el caso de este segundo ítem, se destaca que el 45,45% logra determinar el I.C para la diferencia de medias, pero sólo el 18,18% lo interpreta adecuadamente y, parcialmente 45,45%.

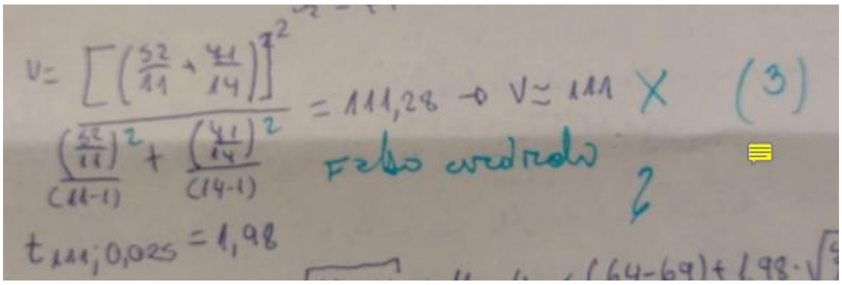
http://doi.org/10.14393/ER-v31e2024-03

Algunos de los procedimientos correctos y errores de los futuros profesores de Matemáticas y Estadística estuvieron en aspectos procedimentales, por ejemplo, asumir varianzas conocidas cuando esto no era explícito, cálculo de la estimación de los grados de libertad para el caso del intervalo de diferencias de medias, errores de procedimiento en la construcción de los IC, búsqueda de cuantiles en la tabla normal estándar o T-Student. Para el caso de los aspectos conceptuales, confunden la varianza muestral con la poblacional, usan el IC para σ^2 conocida, interpretan deterministamente el IC, asumen, sin justificación, que las varianzas son conocidas para el IC de la diferencia de medias y, confunden el nivel de confianza con probabilidad.

Mostramos algunos ejemplos de respuestas de los futuros profesores sobre el cálculo e interpretación de los intervalos de confianza, para el caso de la media poblacional y la diferencia de medias.

Uno de los errores frecuentes que cometieron los futuros profesores fue la estimación de los grados de libertad en el IC para la diferencia de medias, particularmente en el mal uso de la fórmula $\nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left[\left(s_1^2/n_1\right)^2/(n_1-1)\right] + \left[\left(s_2^2/n_2\right)^2/(n_2-1)\right]}$, como se ve en la Figura 1

FIGURA 1: Error en el cálculo de la estimación de los grados de libertad



Fuente: respuesta estudiante 6.

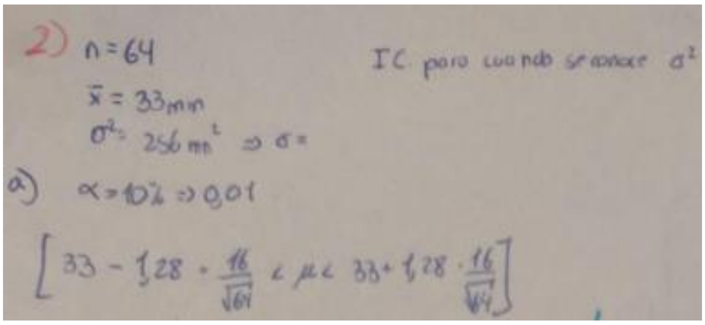
En este caso, vemos que el error procedimental deriva de no considerar los cuadrados en los cocientes $(s_1^2/n_1)^2$ y $(s_2^2/n_2)^2$ que se ubican en el denominador de la fórmula, tal vez, al pensar que ya se trataba de la varianza, de acuerdo con los datos entregados en la actividad. Este error,

http://doi.org/10.14393/ER-v31e2024-03

lleva a los futuros profesores a obtener 111 grados de libertad y, consecuentemente, no determinar el cuantil correcto en la tabla de la distribución T-Student con un 95% de confianza.

Otro de los errores cometidos por los futuros profesores es la determinación del valor crítico en la construcción del IC para la media cuando se desconoce la varianza poblacional (Figura 2).

FIGURA 2: Error en la identificación del cuantil



2) $n = 64$ IC para cuando se conoce σ^2

$\bar{x} = 33 min$

$\sigma^2 = 256 mn^2 \Rightarrow \sigma =$

a) $\alpha = 10\% \Rightarrow 0,01$

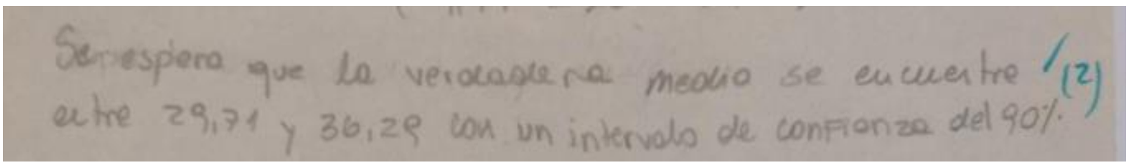
$\left[33 - 1,28 \cdot \frac{16}{\sqrt{64}} < \mu < 33 + 1,28 \cdot \frac{16}{\sqrt{64}}\right]$

Fuente: respuesta estudiante 1.

En este caso, el estudiante 1 asume que la varianza es conocida. Con base en esto, el cuantil lo busca en la tabla de la distribución normal estándar, pero confunde que el IC pedido es bilateral, lo que hace necesario que el cuantil Z_α sea dividida por cada una de las dos colas, cubriendo el área α, para los cuantiles $Z_{\alpha/2} = 1,64$ inferior y superior.

En la interpretación de los IC, como se muestra en la Tabla 1 y Tabla 2, no hubo futuros profesores que interpretaran correctamente los IC construidos. Para el caso de la interpretación del IC para la media, se presenta una respuesta parcialmente lograda (Figura 3), que, en este caso, el profesor asumió varianza conocida, que fue lo que se repitió en el resto de los futuros profesores.

FIGURA 3: Interpretación parcialmente lograda del IC para la media



Se espera que la verdadera media se encuentre entre 29,71 y 36,29 con un intervalo de confianza del 90%. /(2)

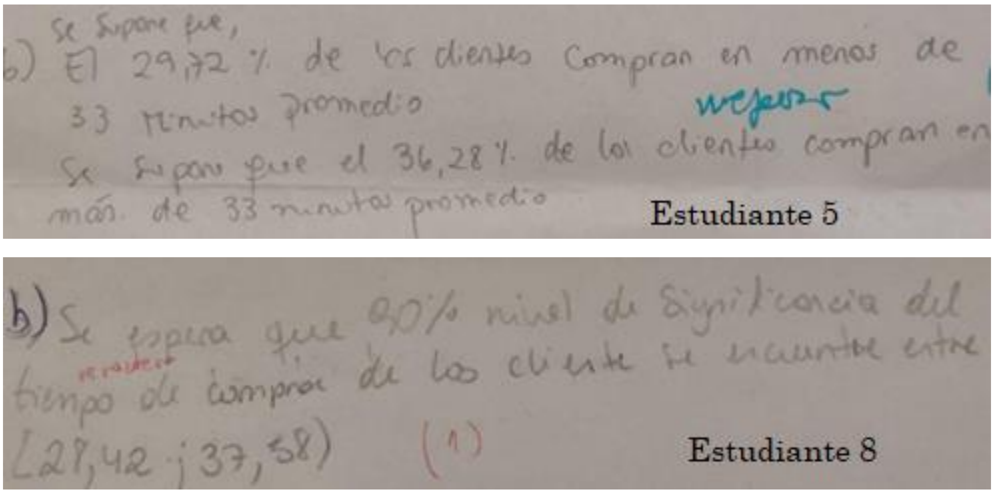
Fuente: respuesta estudiante 2.

http://doi.org/10.14393/ER-v31e2024-03

En este caso, podemos observar que la estudiante realiza una interpretación medianamente correcta, dado que identifica la posibilidad que el intervalo de confianza, con base en una muestra, contenga a la verdadera media poblacional. No obstante esto, no es claro lo que quiere decir al mencionar “con un intervalo de confianza del 90%”, pues podría estar confundiendo el nivel de confianza con el área que se corresponde a las colas. Lo anterior, considerando que la profesora en formación asuma que la distribución es normal.

En este mismo sentido, encontramos casos donde los futuros profesores entregan una interpretación determinista de los IC de confianza, tanto para la estimación de la media μ, como para la diferencia de medias $\mu_1 - \mu_2$ (Figura 4).

FIGURA 4: Interpretación determinista del intervalo de confianza en ambos casos



Fuente: respuestas estudiantes 5 y 8.

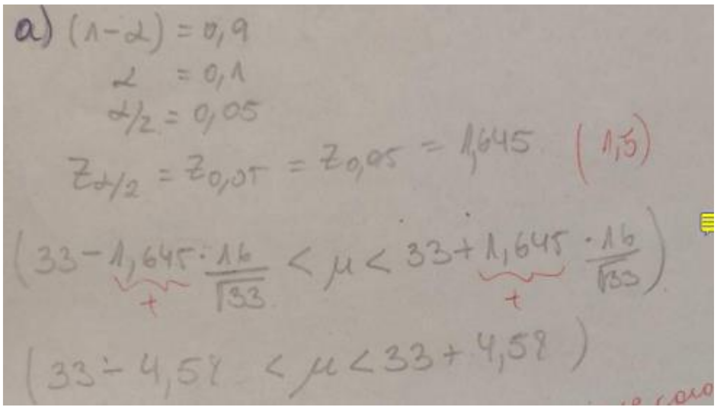
Se puede observar que la futura profesora (estudiante 5) interpreta el IC en términos de la media muestral y no la poblacional. Por otro lado, realiza una interpretación considerando como porcentajes, tanto el límite inferior y superior del IC. Considera las regiones que no incluyen a la verdadera media poblacional, es decir los valores de la media muestral menores que $-Z_{\alpha/2}$ y aquellos mayores

http://doi.org/10.14393/ER-v31e2024-03

que $Z_{\alpha/2}$ de la media muestral, no visualizando que la estimación del verdadero valor de la media poblacional está en la región delimitada por el $1 - \alpha$. En el caso del estudiante 8 se puede observar que si bien pone en su respuesta "se espera", cree erróneamente que el 90% representa el nivel de significancia. Además de ello, no plantea que el tiempo se debe referir al verdadero tiempo promedio de la población, y no considera la aleatoriedad de los intervalos en el muestreo, pues asegura que este tiempo promedio se va a encontrar (deterministamente) entre 28,42 y 37,58.

Otra confusión detectada (estudiante 8) es que, en la construcción del IC para la media con varianza desconocida, el reemplazar en la fórmula $\bar{X} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$, en vez de anotar $n = 64$, evalúa para $n = 33$ (Figura 5).

FIGURA 5: Error de procedimiento al confundir valor del tamaño de la muestra con el valor de la media

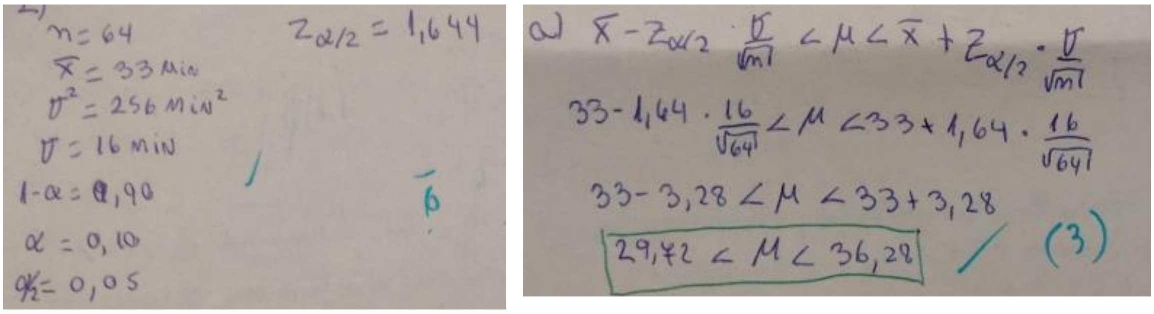


Fuente: respuesta estudiante 8.

Este error al reemplazar $n = 33$ en la fórmula del IC de confianza se puede deber a un despiste, dado que el estudiante anota correctamente los datos que se desprenden del enunciado de la actividad. Por ejemplo, ubica correctamente el valor del cuantil Z (aun cuando considera varianza conocida), cuando esto podría representar mayor complejidad y el uso de la distribución normal podría justificarse a partir del TLC.

Un aspecto interesante en los procedimientos de los futuros profesores es que 10 de ellos consideran que la varianza es conocida $\sigma^2 = 256\ min.$, aun cuando la media $\bar{x} = 33\ min.$ y la varianza $s^2 = 256\ min.$ vienen dadas de manera explícita para una muestra de tamaño $n = 64$, (Figura 6).

FIGURA 6: Procedimiento donde se asume que la varianza es conocida

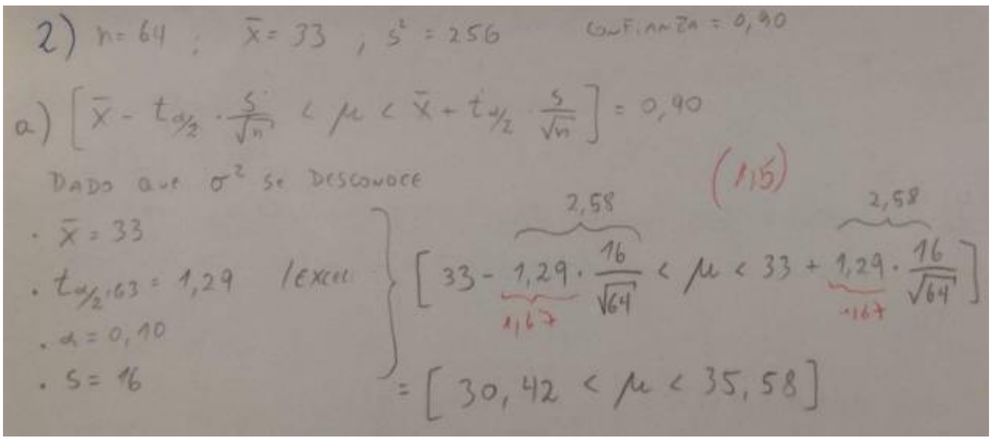


Fuente: respuesta estudiante 6.

Aun cuando esta consideración se puede ver como un error conceptual por parte de los estudiantes al no considerar la estimación de la varianza y, consecuentemente el uso de la distribución T-Student, los procedimientos identificados, considerando la distribución normal, son correctos, lo que representa un 82,82% de respuestas que se logran parcialmente al responder el ítem 1a.

Un caso contario, se da con futuro profesor que no asume la normalidad, es decir, identifica que $s^2 = 256$ es una estimación de la varianza poblacional (Figura 7).

FIGURA 7: Procedimiento correcto con varianza desconocida, pero no determina el percentil asociado



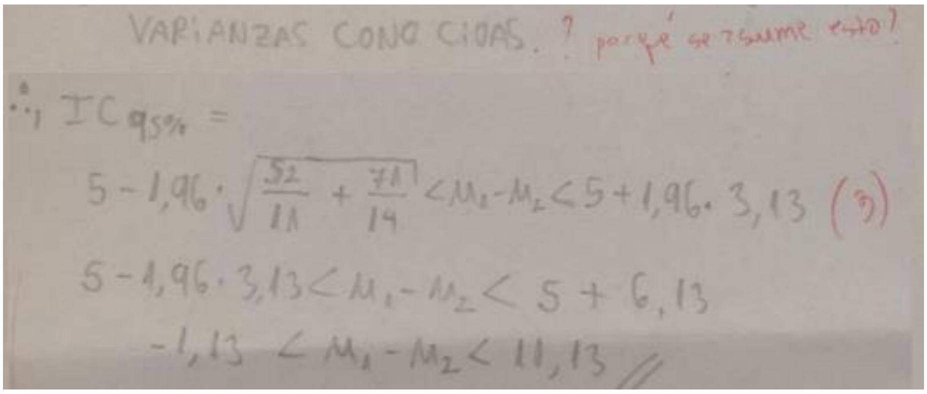
Fuente: respuesta estudiante 6.

http://doi.org/10.14393/ER-v31e2024-03

Este futuro profesor, si bien interpreta adecuadamente el contexto del problema, identificando que $s^2 = 256$ es una estimación de la varianza poblacional (Figura 7) y usando la distribución T-Student, no considera que el IC es bilateral y ubica el valor del percentil $t_{\alpha,n-1} = t_{0,1,63} = 1{,}29$ y no el percentil $t_{\alpha/_{2};\,n-1} = t_{0{,}05\,;\,63} = 1{,}669$.

Finalmente, nos encontramos con el caso del IC para la diferencia de medias con varianzas desconocidas y desiguales. Un estudiante considera, sin mayor suposición, varianzas desconocidas pero iguales, lo que supone utilizar una estimación de la varianza S_p^2 bajo la distribución T-Student, en la que los grados se libertad se determinan a partir de la expresión $n_1 + n_2 - 2$ (donde n_1 y n_2 son los tamos de cada muestra) de las dos muestras independientes (Figura 8).

FIGURA 8: Construcción del intervalo de confianza asumiendo varianzas conocidas



Fuente: respuesta estudiante 11.

Se puede observar que este futuro profesor, en primer lugar, calcula erróneamente la diferencia de las medias $\mu_1 - \mu_2 = 5$ cuando la diferencia correcta es $\mu_1 - \mu_2 = -5$, en segundo lugar, sólo utiliza la expresión $\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ del denominador del pivote y no realiza el cálculo de la varianza estimada ponderada para S_p^2, lo que evidencia una dificultad en la comprensión de fenómenos donde se puede o no asumir ciertas condiciones iniciales para que la estimación, en este caso, por intervalos, sea más robusta al verdadero parámetro poblacional.

http://doi.org/10.14393/ER-v31e2024-03

Discusión y conclusiones

Dado el objetivo de esta investigación "evaluar la comprensión que tienen futuros profesores de Matemática y Estadística sobre los intervalos de confianza" este trabajo aporta información exploratoria sobre la forma en que futuros profesores de Matemática y Estadística comprenden los intervalos de confianza en relación con la construcción, el uso de las definiciones, los supuestos y las propiedades que subyacen. Si bien hay diversos trabajos que han aportado información sobre este tópico (e.g. Andrade; Fernández, 2016; Henriques, 2016; Reaburn, 2014; Roldán et al. 2020) en diferentes contextos profesionales, este trabajo aporta en la formación de profesores de Matemática que enseñarán estadística y que, al menos, en el contexto chileno, es un campo fértil de investigación. En este sentido, los resultados reportados en esta investigación se muestran coherentes con los nuevos estándares de formación inicial docente (MINEDUC, 2021b), los que demandan que los futuros profesores de Matemática sean capaces de conducir la enseñanza de la estadística y la probabilidad, siendo los intervalos de confianza un tema a tratar en 3° y 4° medio (16 a 18 años), y que justifica la evaluación en la comprensión de futuros profesores.

Como parte de los resultados encontrados se evidencia que los futuros profesores de Matemática y Estadística muestran diversas dificultades en la comprensión, tanto conceptual, como procedimental, sobre los intervalos de confianza. En general, muestran errores en la estimación de los grados de libertad, se obvian los supuestos en la construcción de los intervalos de confianza, que son resultados similares a los reportados en otros trabajos (Behar, 2001; Coulson et al., 2010; Henriques, 2016), pero en áreas disciplinares diferentes a la formación del profesorado.

La mayor dificultad evidenciada en los futuros profesores es considerar que la fórmula de la varianza poblaciones es indistinta a la varianza de la muestra, aun cuando las actividades explicitan el tamaño de la muestra. Esto los lleva a usar la distribución normal y no la T- Student (caso que podría justificarse bajo el TLC).

Como parte de los errores conceptuales, así mismo, en algunos futuros profesores de Matemática y Estadística, emergieron interpretaciones deterministas, aspecto reportado en Roldán et al. (2020), interpretaciones que pueden estar relacionadas con la forma de pensamiento determinista y no estadístico. En este sentido, los futuros profesores no consideran que el $(1-\alpha)\%$ determina la probabilidad en que el I.C contendrá al verdadero parámetro poblacional (Cumming, William, Fidler; 2004). Además, se encuentran interpretaciones de los intervalos de confianza como porcentajes, de decir, no se considera el tipo de variable de estudio (actividad 1) y se asume que los limites superior e inferior son porcentajes donde debe caer el verdadero parámetro.

Los errores procedimentales, están en la línea de lo reportado por Olivo, Batanero y Díaz (2008), Rashidah et al. (2018) y Roldán et al. (2020), es decir, calculan erróneamente el cuantil derivado de la tabla normal estándar o T-Student, la estimación de los grados de libertad en el cálculo del IC para la diferencia de medias con varianzas desconocidas, pero desiguales, no consideran que los IC son bilaterales y no dividen la probabilidad α para ambas colas de la distribución, errores de cálculo en la evaluación de la fórmula de los intervalos de confianza para la media μ o la diferencia de medias $\mu_1 - \mu_2$ y, confundir el tamaño de la muestra con el valor de la media muestral al reemplazar en la fórmula, como también, se pasan por alto relaciones relevantes entre los elementos que son parte de los intervalos de confianza, lo que ha sido reportado en investigaciones como la de Andrade; Fernández (2016); Crooks et al. (2019); Roldán et al. (2020) etc.

Tanto los errores procedimentales como los conceptuales en la construcción de los intervalos y las interpretaciones que de estos se derivan plantea la necesidad de revisar, tanto los programas de estudio (MINEDUC, 2021a) como los Estándares de formación inicial docente (MINEDUC, 2021b), pues estos futuros profesores de Matemáticas cometen errores y tienen interpretaciones similares a los de otras áreas de conocimiento, pero aquí, la dificultad radica en que estos conocimientos "sesgados" serán posteriormente enseñados en aula de secundaria, transmitiéndose [probablemente], los mismos errores y formas de interpretar IC.

Esta primera aproximación exploratoria, a partir de los hallazgos encontrados, abre un espacio relevante en el contexto de la formación de futuros profesores de Matemática [y estadística], dado que, si bien las investigaciones han mostrado las dificultades asociadas a la comprensión de los IC en diversos contextos, en el profesorado en formación los resultados son limitados, y como se mencionó, estos resultados si bien incluyen un grupo reducido de futuros profesores de una Universidad formadora de profesores, deja en evidencia algunos de los aspectos que podrían ser considerados por formadores de profesores. Aun cuando varios de los resultados encontrados en esta investigación, son similares a los de otros trabajos en área diferentes a la del profesorado, el hecho que sea en futuros profesores de Matemática lo hace especialmente delicado, considerando que hoy en día el objetivo está en el desarrollo de ciudadanos estadísticamente alfabetizados (Gal, 2002).

En este último caso, trabajos en la línea de la formación de profesores de Matemáticas, son especialmente relevantes para profundizar en la comprensión que tienen y van desarrollando los futuros profesores de Matemática, pues como se ha reportado en diversas investigaciones en contexto curricular de formación de profesores (e.g. Bakker; Derry, 2011; Del Pino; Estrella, 2012; Sánchez; Ruiz, 2022), la enseñanza de la Estadística se ha mostrado atomizada y con un foco en el trabajo mecanicista, dejando de lado la promoción de los aspectos conceptuales que subyacen a la disciplina.

Algunas de las recomendaciones que se pueden derivar de este tipo de trabajo van en dos direcciones: (1) transitar hacia una enseñanza de la estadística considerando investigaciones novedosas a la luz de las propuestas curriculares de enseñanza y aprendizaje, por ejemplo, el proyecto GAISE (Franklin et al, 2007), promover un ambiente de razonamiento estadístico (Garfield; Ben-zvi, 2009) y, (2) centrar la atención en el formador de formadores, pues no está demás cuestionarse que uno de los factores que puede generar una enseñanza mecanicista y atomizada (como se ha reportado), pueda deberse a la forma en cómo fueron enseñados estos profesores, es decir, por los docentes que enseñan estadística en cursos universitarios de formación de profesores

Esto último, más que ser una crítica a la luz de los resultados derivados de esta investigación, pretende sentar y abrir la reflexión en el formador de formadores, pero especialmente, en aquel que está a cargo de la formación de los futuros profesores de Matemática, que, en particularmente, enseñarán estadística y deben ser capaces de estar en la línea de las actuales demandas que requiere el sistema educativo y la sociedad.

Referencias

ANDRADE, L.; FERNÁNDEZ, F. Interpretation of Confidence Interval Facing the Conflict. *Universal Journal of Educational Research*, v. 4, p, 2687–2700, 2016. DOI: http://dx.doi.org/10.13189/ujer.2016.041201.

BAKKER, A.; DERRY, J. Lessons from inferentialism for statistics education. *Mathematical Thinking and Learning*, v. 13, n. 1–2, p. 5–26, 2011. DOI: http://dx.doi.org/10.1080/10986065.2011.538293.

BEHAR, R. *Aportaciones para la mejora del proceso de enseñanza-aprendizaje de la estadística*. Trabajo de Tesis Doctoral. Universidad Politécnica de Cataluña, 2011.

COULSON, M., HEALEY, M.; FIDLER, F.; CUMMING, G. Confidence intervals permit, but do not guarantee, better inference than statistical significance testing. *Frontiers in Psychology*, v. 1, n. 26, p. 1-9, 2010. DOI: https://doi.org/10.3389/fpsyg.2010.00026.

CROOKS, N. M.; BARTEL, A. N.; ALIBALI, M. W. Conceptual knowledge of confidence intervals in psychology undergraduate and graduate students. *Statistics Education Research Journal*, v. 18, n. 1, p. 46–62, 2019. DOI: https://doi.org/10.52041/serj.v18i1.149.

CUMMING, G.; WILLIAMS, J.; FIDLER, F. (2004). Replication, and researchers' understanding of confidence intervals and standard error bars. *Understanding Statistics,* v. 18, n. 3*,* p. 299-311, 2004.

DEL PINO, G.; ESTRELLA, S. Educación estadística: relaciones con la matemática. *Pensamiento Educativo, Revista De Investigación Latinoamericana (PEL)*, v. 49, n. 1, p. 53–64, 2012. DOI: https://doi.org/10.7764/PEL.49.1.2012.5.

http://doi.org/10.14393/ER-v31e2024-03

ENGEL, J. On teaching bootstrap confidence intervals. In C. READING (Ed.), *Proceedings of the eighth International Conference on Teaching Statistics*. Voorburg, The Netherlands: International Statistical Institute. 2010. Disponible en: http://www.stat.auckland.ac.nz/~iase/publications.php.

FRANKLIN, C.; KADER, G.; MEWBORN, D.; MORENO, J.; PECK, R.; PERRY, M.; OTROS. *Guidelines for assessment and instruction in statistics education (GAISE) report: A preK-12 curriculum framework*. Alexandria, VA: American Statistical Association, 2007.

GAL, I. Adults' Statistical literacy: meanings, components, responsibilities. *International Statistical Review*, Nederlandn, n. 70, p. 1-25, 2002.

GARFIELD, J. B.; BEN-ZVI, D. Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. *Teaching Statistics,* v. 31, n. 3, p. 72–77, 2009.

HENRIQUES, A. Students' difficulties in understanding of confidence intervals. En D. BEN-ZVI.; K. MAKAR (Eds.), *The Teaching and Learning of Statistics* (pp. 129-138). Cham, Switzerland: Springer, 2016.

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

HOEKSTRA, R.; MOREY, R. D.; ROUDER, J. N.; WAGENMAKERS, E. J. Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, v. 21, p. 1157–1164, 2014.

MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE, MECD. *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato.* Madrid: Autor, 2015.

MINEDUC. *Bases curriculares 2012. Matemática, educación básica*. Santiago: Ministerio de Educación, 2012.

MINEDUC. *Programa de Estudio Probabilidades y Estadística Descriptiva e Inferencial 3° y 4° medio – Formación Diferenciada*. Unidad de Currículum y Evaluación Ministerio de Educación, 2021a.

MINEDUC. *Estándares para la formación inicial docente de Matemáticas*. Unidad de Currículum y Evaluación Ministerio de Educación, 2021b.

NCTM. *Principles and standards for school mathematics*. Reston, VA: NCTM, 2000.

OLIVO, E.; BATANERO, C. Un estudio exploratorio de dificultades de comprensión del intervalo de confianza. *Union*, v. 12, p. 37–51, 2007.

OLIVO, E.; BATANERO, C., DÍAZ, C. Dificultades de comprensión del intervalo de confianza en estudiantes universitarios. *Educación Matemática*, v. 20, n. 3, p. 5-32, 2008.

PÉREZ-SERRANO, G. Investigación cualitativa. Retos e interrogantes. Madrid: La Muralla, 1994.

http://doi.org/10.14393/ER-v31e2024-03

RASHIDAH, N.; RAZAK, F. A.; BAHARUN, N.; ARUL, E. S. G. Investigating Students' Difficulties in Understanding Confidence Intervals in Linear Regression Models. *International Journal of Engineering & Technology*, v. 7, n. 4.33, p. 60-64, 2018. DOI: http://dx.doi.org/10.14419/ijet.v7i4.33.23485.

REABURN, R. Students' understanding of confidence intervals. In K. MAKAR, B. DE SOUSA.; R. GOULD (Eds.), *Proceedings of the Ninth International Conference on Teaching Statistics*. Voorburg, The Netherlands: IASE, 2014.

ROLDÁN, A. F.; BATANERO, C.; ÁLVAREZ-ARROYO, R. Comprensión del intervalo de confianza por estudiantes de Bachillerato *Avances De Investigación En Educación Matemática*, v. 18, p. 103–117, 2020. DOI: https://doi.org/10.35763/aiem.v0i18.284.

SÁNCHEZ, N.; RUIZ, B. Análisis de las actividades que proponen dos libros de texto de educación primaria. Un acercamiento comparativo desde la perspectiva de la inferencia informal. *REXE- Revista De Estudios Y Experiencias En Educación*, v. 21, n. 46, p. 76–101, 2022. DOI: https://doi.org/10.21703/0718-5162.v21.n46.2022.004.

SHAUGHNESSY, J. M. Research on statistics learning and reasoning. In F. K. LESTER JR. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957–1009). Greenwich, CT: Information Age/NCTM, 2007.

STAKE, R. *Investigación con estudio de casos*. 4. Ed. Madrid: Morata, S.L, 2007.

PÉREZ-SERRANO, G. *Investigación cualitativa: retos e interrogantes*. Madrid: La Muralla, 1994.

VÁSQUEZ, C.; Y ALSINA, Á. Enseñanza de la Probabilidad en educación primaria. Un desafío para la formación inicial y continua del profesorado. *Números. Revista de Didáctica de las Matemáticas*, v. 85, p. 5-23, 2014. Disponible en: http://www.sinewton.org/numeros/numeros/85/Articulos_01.pdf.

YÁÑEZ, G.; BEHAR, R. Interpretaciones erradas del nivel de confianza en los intervalos de confianza y algunas explicaciones plausibles. En M. J. González, M. T. González y J. Murillo (Eds.). *Investigación en Educación Matemática XIII, Comunicaciones de los grupos de investigación*. XIII Simposio de la SEIEM. Santander: SEIEM. 2009. Disponible en: https://seiem.es/docs/comunicaciones/ GruposXIII/depc/Yanez_Behar_R.pdf.

YAREMKO, R. M.; HARARI, H.; HARRISON, R. C.; LYNN, E. *Handbook of research and quantitative methods in psychology: For students and professionals.* Hilldale, NJ: Erlbaum, 2013.

WACKERLY, D. D.; MENDENHALL III, W.; SCHEAFFER, R. L. *estadística matemática con aplicaciones*. México, D.F.: Cengage Learning Editores, 2010.

WALPOLE, R.E.; MYERS, R.H.; MYERS, S.L.; YE, K. *probabilidad y estadística para ingeniería y ciencias*. 9° edición. Pearson Educación México, 2012.

Recibido en junio de 2023.

Aprobado en noviembre de 2023.