

EVALUACIÓN DEL CONOCIMIENTO Y ARGUMENTACIÓN ADQUIRIDOS POR FUTUROS PROFESORES DE SECUNDARIA DE MATEMÁTICA SOBRE ÍNDICES DE RESUMEN NUMÉRICO

Francisco Rodríguez-Alveal*

frodriguez@ubiobio.cl
orcid.org/0000-0003-2169-0541

Universidad del Bío-Bío

Danilo Díaz-Levicoy**

dddiaz01@hotmail.com
orcid.org/0000-0001-8371-7899

Universidad de Granada

Ana Maldonado-Fuentes***

amaldonado@ubiobio.cl
orcid.org/0000-0002-9198-5882

Universidad del Bío-Bío

Recibido: 20/02/2018

Aprobado: 11/07/2018

RESUMEN

En este artículo, de naturaleza cualitativa, se analizan las respuestas abiertas de estudiantes de Pedagogía, mención Matemática, sobre las medidas de tendencia central (media aritmética, mediana y moda). El objetivo de esta investigación fue describir los conocimientos y argumentos estadísticos adquiridos por estos alumnos en su itinerario formativo. El estudio estuvo compuesto por 19 futuros profesores que respondieron tres preguntas. La primera consistió en analizar el efecto de datos atípicos sobre la media aritmética, la segunda implicó seleccionar la medida de tendencia central más adecuada para resumir información cualitativa ordinal y, en la tercera, debieron comentar las propiedades numéricas, algebraicas y estadísticas de los índices numéricos mencionados. Los resultados permitieron evidenciar que, en general, los participantes no presentan dificultades procedimentales para el cálculo de la media aritmética, mediana y moda, no obstante los argumentos estadísticos entregados son poco plausibles para justificar la elección de estos índices en la toma de decisiones.

Palabras clave: formación de profesores; didáctica de la estadística; evaluación; medidas de tendencia central.

***Francisco Rodríguez-Alveal.** Magíster en Bioestadística (UCHILE). Académico del Departamento Ciencias de la Educación de la Universidad del Bío-Bío (UBB), Chillán, Chile. **Universidad de adscripción:** Universidad del Bío-Bío.

****Danilo Díaz-Levicoy.** Máster en Didáctica de la Matemática (UGR). Doctorando en Ciencias de la Educación en la Universidad de Granada (UGR), Granada, España. **Universidad de adscripción:** Universidad de Granada.

*****Ana Maldonado-Fuentes.** Magíster en Ciencias de la Educación (PUC). Académica del Departamento Ciencias de la Educación de la Universidad del Bío-Bío (UBB), Chillán, Chile. **Universidad de adscripción:** Universidad del Bío-Bío.

EVALUATION OF KNOWLEDGE AND ARGUMENTATION ACQUIRED BY FUTURE HIGH SCHOOL MATH TEACHERS BASED UPON NUMERICAL RESUME INDEXES

ABSTRACT

In this qualitative study, math education students open answers about central tendency measures (arithmetical media, mediana and mode were analyzed). The objective of the study was to describe knowledge and statistical arguments acquired by the students mentioned above in their career formation. The population of the study was 19 future teachers who answered three questions. First, referred to analyzing the affect of atypical about arithmetical media, the second one implied to choose the most suitable central tendency measure to resuming ordinal qualitative information and, the third one, algebraic and statistical numerical properties about numerical indexes already mentioned, must be commented. Results evidenced that, in general, participants do not show procedimental difficulties for measuring arithmetical media, mediana and mode, however statistical arguments given, are not very suitable to justify election of these indexes in decision-making.

Key words: teachers; formation; didactic of statistics; evaluation; measures of central tendency.

Introducción

En Chile, la formación inicial docente, en los últimos quince años ha sido objeto de diversos estudios centrados en el análisis de las políticas de formación de profesores, en su efectividad para producir conocimiento pedagógico general y de contenido matemático (Ávalos, 2014). Debido a que los egresados de carreras de pedagogía han obtenido bajos resultados en matemática, lo que ha suscitado modificaciones en los programas de estudio y la reflexión en los centros de enseñanza (Estrella, Olfos y Mena-Lorca, 2015).

Además, los estudios han evidenciado en la formación del profesorado la existencia de nudos críticos con relación a la formación disciplinar y didáctica en los planes de estudio (Sotomayor-Echenique, Coloma-Tirapegui, Parodi-Sweis, Ibáñez-Orellana, Cavada-Hrepich y Gysling-Caselli, 2013), así como la escasa articulación de los conocimientos con el conocimiento pedagógico del contenido y práctica (Vaillant, 2007). En este sentido, “existe un diagnóstico extendido que alerta sobre la falta de calidad y pertinencia de la formación inicial y la necesidad de acercarla a las demandas reales de la profesión” (Vezub, 2016, p. 2).

Producto de lo anterior, una de las principales expectativas del Decreto Ley N° 20.903 del Ministerio de Educación (MINEDUC, 2016) es que los futuros profesores de Educación Secundaria tengan, al momento de egresar, un dominio de saberes disciplinares y pedagógicos, consensuados en estándares a nivel nacional (MINEDUC, 2012a); y se proyecten con altos desempeños en su labor docente, ya que una buena formación del estudiante depende en gran medida de los conocimientos del profesor. Del mismo modo, se espera que tal como los futuros profesores y los del sistema escolar estén capacitados para poner en práctica las políticas públicas relacionadas con educación y el mejoramiento de los aprendizajes, en coherencia con las demandas ministeriales y la sociedad chilena en general.

También, estos conocimientos deben estar de acuerdo con las directrices curriculares. En este sentido, un cambio importante ocurrió en Matemática con la inclusión del bloque de Estadística y Probabilidad desde el primer curso de Educación Primaria (MINEDUC, 2012b), modificando los contenidos a enseñar en las aulas y en las instituciones formadoras de profesores. En este bloque se desarrollan habilidades propias de la estadística, modelamiento de datos e interpretación de ellos (MINEDUC, 2009; Rodríguez-Alveal y Sandoval, 2012).

Estudios a nivel internacional (e.g., Arteaga, Batanero, Contreras y Cañadas, 2016; Estrada, Batanero y Fortuny; 2004; Gea, Arteaga y Cañadas, 2017; Jacobbe, 2012; Ortiz, Font y Mayén, 2009) y nacional (e.g., Díaz-Levicoy, Sepúlveda, Vásquez y Opazo, 2016; Estrella, 2016; Estrella *et al.*, 2015; Rodríguez-Alveal y Sandoval, 2012), han evidenciado que los profesores de diferentes niveles de formación presentan habilidades descendidas en temas de estadística. Específicamente, los hallazgos han permitido constatar la necesidad de profundizar en los conocimientos de los estudiantes en su formación docente más allá de lo algorítmico, integrando la comprensión y el razonamiento acerca de las medidas de tendencia central como parte de los índices de resumen numéricos, con el objeto de lograr una alfabetización estadística (Gal, 2002; Rodríguez-Alveal, 2017; Watson, 1997).

En este sentido, el presente artículo es una extensión del estudio realizado por Rodríguez-Alveal, Maldonado-Fuentes y Sandoval (2016), con el objetivo de describir los conocimientos y argumentos estadísticos adquiridos por futuros profesores de Educación Secundaria durante su itinerario formativo, como una forma de contribuir a la toma de decisiones de los formadores de formadores, respecto de cuáles son las dificultades que presentan los futuros docentes en relación con los conocimientos de las medidas de tendencia central.

Fundamentos

Como menciona Hacking (2005, p. 128): “la estadística comenzó siendo el estudio sistemático de hechos cuantitativos acerca del estado”, la cual puede definirse, en términos generales, como *la ciencia de los datos* (Moore, 1991), siendo de interés el razonamiento a partir de los datos empíricos.

Definición, que se alinea con lo establecido en el bloque de Estadística y Probabilidad de las directrices curriculares del Ministerio de Educación chileno (MINEDUC, 2012b). En coherencia, la enseñanza de la estadística pretende, a lo largo de los años de formación obligatoria, que los estudiantes desarrollen las habilidades de leer y analizar información así como hacer inferencias a partir de distribuciones de datos (Rodríguez-Alveal *et al.*, 2016; Shaughnessy, 2007). Para el logro de este propósito, el uso de las medidas de tendencia central resulta relevante, dado que estas juegan un rol fundamental en el resumen numérico de los datos; específicamente las medidas de tendencia central: la media aritmética, la mediana y la moda.

En primer lugar, resulta de interés profundizar en el objeto estadístico de la *media aritmética* que ha estado presente en el currículo escolar por más de cien años (Jacobbe y Carvalho, 2011). Este resumen numérico posee propiedades algebraicas, numéricas y estadísticas. Entre ellas se destaca que es un operador lineal, se ve afectada por observaciones extremas y no tiene la estructura de grupo (Mevarech, 1983). Este objeto estadístico es base para la comprensión de conceptos tales como variables aleatorias, distribuciones de probabilidad y de muestreo, Teorema del Límite Central, entre otros. Al mismo tiempo, a nivel de inferencia estadística, el promedio o media aritmética muestral es un “buen” estimador del promedio poblacional, debido a que satisface las propiedades estadísticas de insesgamiento, consistencia, suficiente y varianza mínima.

Otro antecedente a considerar, es la introducción del concepto de *media aritmética* en las directrices curriculares chilenas a partir de Enseñanza Primaria (MINEDUC, 2012b). En particular, a partir de quinto años de Educación Primaria se define la media aritmética como “el cociente entre la suma de los valores numéricos de la variable y la cantidad total de datos” (Ho, Kee y Ramakrishnan, 2017, p. 299), mientras que el libro de texto de séptimo año se define a la media aritmética como “la cantidad total de la variable distribuida en partes iguales” (Merino, Muñoz, Pérez, Rupin, 2016, p. 313).

En segundo lugar, otro estadístico presente en las directrices curriculares a nivel de Educación Primaria es la *mediana*, la que en el libro de texto de Marín, Castillo y Fernández (2011) de quinto año la definen de la siguiente forma:

Dada una colección de datos cuantitativos, se llama **mediana** al dato central, es decir, al dato que queda en el medio luego de haberlos ordenado en orden creciente o decreciente. Si la cantidad de datos es impar la mediana corresponde a uno de los datos de la colección. Si la cantidad de datos es par no existe un único dato central, sino dos; en este caso la mediana se define como el promedio aritmético de estos dos valores. (p. 148)

De manera similar, en el libro de texto del estudiante de séptimo entregado por el MINEDUC (Merino *et al.*, 2016, p. 321), se hace referencia a que “la mediana corresponde al valor que ocupa el término central de un conjunto de datos una vez ordenados de menor a mayor o viceversa”. En este mismo sentido, estudios como los de Barr (1980) han llegado a la conclusión que alumnos entre 17 y 21 años interpretan la mediana como el centro de *algo*. Por su parte, Cobo y Batanero (2000) sugieren que los estudiantes se encuentran con dificultades para calcular la mediana si parten de las representaciones gráficas de las frecuencias acumuladas, ya que no están acostumbrados a las funciones discontinuas.

Además, otra medida que resulta de interés es la *moda* o valor modal, que a nivel de Educación Primaria, de acuerdo con los libros de texto, es definida como “la variable con mayor frecuencia, y no la frecuencia más grande” (Merino *et al.*, 2016, p. 316), la cual genera dificultades para su comprensión, debido a que alude implícitamente a variables de naturaleza cualitativa, haciendo isomorfa la naturaleza de la variable como símil de información cuantitativa. No obstante lo anterior, en el mismo libro de texto, se

entrega una segunda definición en la cual “se llama moda (Mo) de un conjunto de datos a la variable que presenta mayor tendencia de ocurrencia” (*op. cit.*, p. 317), también, se menciona que la moda no es única y puede no existir.

Hemos incluido las definiciones de estos tres estadísticos en los libros de texto de Educación Primaria, por la importancia de este recurso pedagógico en el proceso de instrucción y por tratarse del primer acercamiento del estudiante con este objeto estadístico, pudiendo generar obstáculos en su aprendizaje.

Finalmente, es importante una adecuada interpretación de los índices de resumen numérico, para obtener conclusiones validas y de calidad. Por ejemplo, la media aritmética se ve afectada por observaciones extremas (datos atípicos); es decir, si la distribución de ellos es asimétrica positiva o negativa. Este hecho podría representar un nudo crítico de la enseñanza dado que en tal caso no sería recomendable considerar la media aritmética como resumen numérico debido a su pérdida de robustez, siendo aconsejable considerar la mediana como un índice más estable en esta situación.

De acuerdo con los antecedentes expuestos, resulta de interés estudiar, de modo descriptivo, los conocimientos y argumentos estadísticos que declaran futuros profesores de Educación Secundaria sobre las medidas de tendencia central.

Metodología

Para dar cumplimiento a nuestro objetivo hemos seguido una metodología de tipo cualitativa, utilizando como método el análisis de contenidos en las respuestas entregadas por los profesores en formación a las preguntas planteadas, del cual se extrajeron codificaciones que permitieron la generación de categorías que permitió evidenciar contrastes y similitudes de los escritos de los participantes. El procesamiento de la información fue realizado en el software *Qualitative Data Analysis* (QDA).

Muestreo y tamaño muestral

Para efectos del estudio se consideró un muestreo intencionado por disponibilidad (McMillan y Schumacher, 2005). La muestra estuvo compuesta por 19 estudiantes de octavo semestre de la carrera de Pedagogía en Educación Matemática de una universidad del centro-sur de Chile. Todos ellos participaron de manera voluntaria, siguiendo el protocolo de confidencialidad de información acorde a las políticas institucionales, como parte de una actividad formativa de la asignatura *Didáctica de la Estadística* en el año académico 2017. La edad promedio de los estudiantes considerados en el estudio fue de 21,5 años ($D.E= 1,2$ años), con una distribución asimétrica positiva ($C.A=0,86$), situación que evidencia que en general los encuestados se pueden clasificar como adulto joven. En relación al género, el grupo es mayoritariamente femenino (52,6%). Asimismo, dentro de las características académicas, se destaca que provienen en general de colegios de modalidad particular subvencionado (42,1%) y colegios municipalizados (26,3%).

Desde el punto de vista socioeconómico, y considerando que el país cuenta con la distribución en quintiles como factor fundamental a la hora de calificar para ayuda estatal en el caso de la población estudiantil en Educación Superior. De acuerdo a ello, se trata de un grupo que en general presenta un bajo nivel socioeconómico, lo cual es ratificado en que el 52,6% estudia bajo la modalidad de gratuidad, recibiendo aporte estatal para el pago de matrícula y/o arancel en sus estudios.

Instrumento

Para efectos del estudio se aplicó un instrumento de lápiz y papel con tres preguntas de desarrollo, las que se eligieron por evaluar diferentes aspectos de las medidas de tendencia central.

La primera pregunta (Figura 1) fue tomada de Estrella (2016) y de Garret y García (2008), la que fue creada a partir de un problema de respuestas de elección múltiple desarrollado por Garfield y Konold (1992). Es una actividad de Estadística Descriptiva cuyo objetivo es evaluar el uso de la media aritmética como mejor estimador de un conjunto de datos en presencia de variabilidad producto de errores de medida, la influencia de los valores atípicos en el cálculo de la media aritmética; la confusión entre el uso de los índices de resumen numérico, así como la importancia del contexto de los datos.

Pregunta 1: A petición del profesor de Educación Física, 10 alumnos registraron (en forma independiente y simultánea) el tiempo recorrido por un estudiante en la distancia de los 100 m. Los tiempos registrados (en segundos) fueron los siguientes:

15,05 14,95 15,05 15 10 15 14,9 15 14,95 15

¿Qué tiempo debe considerar el profesor como estimación representativa del tiempo real recorrido por el grupo de estudiantes y por qué?

Figura 1. Primera pregunta del cuestionario

La ventaja de incorporar este ítem radica en que discrimina entre el uso algorítmico de la media aritmética y la comprensión relacional del concepto, dado que el futuro profesor debe considerar la importancia del contexto en el cual han sido tomados los datos (Estrella, 2016). Además, pone en juego la propiedad de que *la media aritmética es un estadístico poco robusto, sensible a los valores atípicos* (Konold y Pollatsek, 2004).

La segunda pregunta (Figura 2) fue tomada de Godino (1999). La importancia de ítem radica en que el futuro profesor pueda identificar el índice de resumen numérico más adecuada para datos de naturaleza cualitativa ordinal. La ventaja de incorporar este ítem es ir más allá de la media aritmética, considerando la naturaleza y el contexto de los datos.

Pregunta 2: Un profesor califica a sus alumnos del siguiente modo: I: Insuficiente, A: Aprobado, N: Notable, S: Sobresaliente. A continuación se entregan las notas de ambos grupos de alumnos:

Grupo 1: I A A N N S S I I I A A A N S S I A A S S S S

Grupo 2: S S I I A N A N I I S N A S I N N A.

- a) ¿Qué grupo ha obtenido mejores notas?
b) ¿Cuál sería la medida de tendencia central más apropiada para representar estos datos? Justificar brevemente.
c) ¿Cuál es el grupo que presenta mejores notas?

Figura 2. Segunda pregunta del cuestionario

La tercera pregunta (Figura 3) estaba asociada a indicar las propiedades numéricas, algebraicas y estadísticas de las medidas de tendencia central (media aritmética, mediana y moda). El objetivo de incluir esta actividad es conocer las características teóricas que asocian los futuros profesores a estos resúmenes numéricos. Conocimientos básicos para comprender otros temas, por ejemplo, los de dispersión.

Pregunta 3: *Entregar al menos dos propiedades numéricas, algebraicas o estadísticas de la media aritmética, mediana y moda.*

Figura 3. Tercera pregunta del cuestionario

Resultados

Los resultados se presentan siguiendo el orden de las preguntas en el instrumento de evaluación, incorporando en forma conjunta los resultados del análisis de contenido en tablas de frecuencias y ejemplos de las respuestas aportados por los futuros profesores.

Índices de resumen numérico en presencia de datos atípicos

En la Tabla 1 se explicitan los estadísticos propuestos y la justificación entregada por los estudiantes de Pedagogía en Educación Matemática respecto al resumen numérico de la distribución de un conjunto de datos en presencia de datos atípicos (pregunta 1).

Tabla 1. Distribución porcentual del argumento y justificación de los futuros profesores para decidir sobre el estadístico de resumen en presencia de datos atípicos

Estadístico	Justificación	Porcentaje (n=19)
La media aritmética considerando todos los datos:		63,2
	Mejor representante	36,8
	Fácil de calcular	5,3
	Representa el tiempo real	5,3
	Menor variabilidad entre los datos	5,3
	Sin justificación	10,5
La media sin dato atípico:		10,5
	Dato atípico según la distribución normal	5,3
	Deciden visualmente que el dato es atípico	5,3
El valor más frecuente:		21,1
	Debido a la asimetría de los datos	5,3
	Mejor representante	5,3
	Reduce el error	5,3
	Porque se repite más	5,3
El valor central de la serie ordenada:		5,3

Se observa que el 63,2% de los futuros profesores calculan la media aritmética considerando todas las observaciones, dentro de las argumentaciones dadas por los participantes el 36,8% lo justifica haciendo notar que *es el mejor representante* de los datos. Otras justificaciones entregadas son: *fácil de calcular, representa el tiempo real y menor variabilidad de los datos*. En este caso particular, la alusión a la media aritmética como un *representante* era una respuesta esperada, pues cualquier medida de tendencia central, forma o variabilidad es un *representante* de un conjunto de datos, los cuales están asociados a la distribución y naturaleza de los datos. Por tanto, en algunos casos hay que tomar decisiones respecto de ellos en relación a su robustez y ser capaz de seleccionar el índice de resumen numérico que dado el contexto y la distribución de los datos sea el más adecuado.

En este sentido, habría sido una mejor respuesta aquella que se estudie la presencia de datos atípicos en la distribución de los datos. Sin embargo, esto no sucedió, dado que estos temas mayoritariamente no son estudiados de manera explícita en la asignatura de Estadística en cursos de formación del profesorado, hipótesis que podría ser coherente con el énfasis en la resolución meramente algorítmica de los ejercicios por parte de los sujetos.

Asimismo, se evidencia que el 10,5% de los estudiantes calculan la media aritmética sin considerar el valor 10,0; el cual es clasificado por uno de estos estudiantes como un dato atípico, argumentando al respecto:

Debe considerar el promedio de los 9 datos restantes excluyendo el 10,0 pues no es un valor cercano al grupo curso, un valor atípico de hecho pongo en duda que sea real pues los atletas profesionales rondan ese tiempo como para que un estudiante lo logre, el tiempo representativo de los recorrido es de 14,97 segundos (Estudiante 9).

Y el otro estudiante asume que dicho valor se encuentra en el extremo izquierdo de la cola de la distribución normal. Situación que es justificada como:

El profesor debe considerar como estimación representativa del grupo el valor de 14,988 que es el promedio de los datos excluyendo el tiempo de 10 segundos, ya que es atípico, si lo graficamos en una distribución normal este estaría en un extremo, lo cual indica que no es representativo del grupo (Estudiante 3).

El 5,3% asocia la media aritmética a un *representante del tiempo real*, entre los argumentos dados se encuentra:

La estimación más adecuada para representar el tiempo real recorrido por los estudiantes es el promedio, en este caso 14,49 ya que es un valor que se acerca a lo suficiente a cada uno de los datos sin descartar ninguno de los caso (Estudiante 4).

Estos futuros profesores no realizan ninguna prueba estadística que permita decidir si en el conjunto de valores entregados hay presencia de algún dato atípico (o *outliers* en inglés) cuya identificación sería posible realizar haciendo uso de la medida de dispersión denominada *distancia intercuantil* la cual es sensible a dichos valores. Es decir, valores que se encuentren fuera del intervalo $[Q_1 - 1,5(Q_3 - Q_1); Q_3 + 1,5(Q_3 - Q_1)]$ que en nuestro caso es el valor 10,0 debido a que se encuentra fuera de los límites del intervalo [14,825 y 15,125].

Esta misma situación también puede ser visualizada en el diagrama de cajas que se introduce en el octavo año de Educación Primaria (13-14 años), y que será utilizado en los siguientes cursos de Educación Secundaria, siendo conocido a nivel escolar como *diagrama de cajón*, debido a que consiste en un rectángulo en el cual se pueden identificar los cuartiles. Lo interesante para la enseñanza sería constatar que es posible visualizar características de la distribución de los datos, tales como forma, variabilidad, tendencia central y valores atípicos en esta representación.

Sin embargo, el tratamiento de estas características de los datos suele ser desconocido por los profesores en formación y del sistema escolar, debido a que este concepto en forma específica no es tratado en las asignaturas de Estadística a nivel de pregrado en la formación universitaria, pese a estar presente en el currículo escolar. Otro estadígrafo sugerido por los profesores en formación para resumir numéricamente la información es la moda (21,1%). Entre las argumentaciones dadas se encuentran:

El profesor debe considerar la moda de estos tiempos como estimación representativa, pues es el mejor estimador o representante de estos tiempos, considerando que son asimétricas por lo cual sería 15,0 (Estudiante 8).

El mejor representante sería la moda, ya que los extremos 10 y 15.05 indican que un alumnos pero el cronometro muy ante y a el otro un poco después, la media no solo se tiene un sujeto del estudio y el tiempo representativo no hubiese sido el más óptimo, mientras que con la moda se "reduce" el error al momento de frenar o parar el cronometro siendo este (Estudiante 10).

Dado que la distribución de los datos es un factor a tener en cuenta al momento de elegir el *mejor representante numérico*. La teoría explicita que en el caso de distribuciones asimétricas es recomendable hacer uso de la mediana, ya que no es afectada por las observaciones extremas, lo que sí ocurre con la media aritmética. En este sentido, el 5,3% de los estudiantes de Pedagogía en Matemática proponen a la mediana como el resumen numérico más adecuado, mencionando:

El profesor debería usar la mediana como representante para la estimación del tiempo real recorrido por el grupo de estudiantes porque existe un dato atípico que corresponde al 10,0 el cual perturbaría la estimación si se usara la media como representante de estimación (Estudiante 11).

Índices de resumen para un conjunto de datos ordinales

Una de las exigencias del sistema educacional chileno a la formación de profesores de matemática es que durante su formación adquieran una fuerte base teórica de los contenidos que deben enseñar a nivel de Educación Primaria o Secundaria (Rodríguez-Alveal *et al.*, 2016). En este sentido, la segunda pregunta realizada a los futuros profesores está relacionada con la toma de decisiones en presencia de información de naturaleza cualitativa.

Tabla 2. Distribución porcentual del argumento y justificación de los futuros profesores para decidir sobre el estadístico de resumen en presencia de información cualitativa ordinal

Estadístico	Justificación	Porcentaje (n=19)
La media aritmética		21,1
	Se ve con más claridad el valor	5,3
	Se puede utilizar para datos cualitativos	5,3
	Porque los datos no se distribuyen normalmente	5,3
	Nos da la nota final por grupo	5,3
Mediana		26,3
	Es un eje de equilibrio	5,3
	Representa las dos mitades	5,3
	Numero de observaciones desiguales	5,3
	No entregan argumentación	10,5
Moda		42,1
	Datos cualitativos	15,8
	Permite ver el dato que más se repite	15,8
	El mejor representante	5,3
	No entregan argumentación	5,3
Regla de tres		5,3

En la Tabla 2 se muestran los estadísticos que proponen los estudiantes y la justificación de los mismos en relación a dar respuesta a la interrogante *¿Cuál es la medida de tendencia central más apropiada para resumir información cualitativa ordinal?*

Tal como se observa en el resumen anterior, hay un 21,1% de los futuros profesores que proponen la media aritmética como medida de tendencia central (índice de resumen numérico) en este caso, es decir, asocian propiedades numéricas a un código cualitativo. Algunos de los argumentos son los siguientes:

La medida más apropiada es la media ya que se ve con claridad donde se encuentra el valor del grupo curso y que calificaciones son mayoritarias (Estudiante 3).

El promedio sería la más apropiada puesto que nos dará la nota final por grupo, permitiéndonos compararlos y ver cuál es mejor (Estudiante 6).

La mejor sería la media ya que si observamos la mediana en el grupo 1 sería A, grupo 2 tendría 2 que sería A y N, si vemos la moda el grupo 1 sería S, mientras

que el grupo 2 sería I y N, elementos poco representativos pero si observamos la media en el grupo 1 sería 6, que está en A, mientras que en el grupo 2 sería 4 que sería I, que sería un representante más comparable (Estudiante 10).

La media, ya que no se distribuyen de forma normal (Estudiante 13).

Estos argumentos evidencian la falta de conocimientos teóricos, ya que la media aritmética es una medida de tendencia central que resume información cuantitativa y su cálculo no está condicionado a que los datos se distribuyan de manera normal.

Por otro lado, se muestra que un 26,3% de los estudiantes para profesores proponen a la mediana como la medida de tendencia central más adecuada para el tipo de información. A continuación se presentan algunos argumentos entregados por los estudiantes:

La mejor medida sería la mediana ya que la cantidad de datos en los grupos es desigual y la moda no es representativa en el campo problemático (Estudiante 17).

La mediana porque como las notas están representadas como letras y la cantidad de alumnos en los grupos no es la misma considero que es mejor utilizarla (Estudiante 1).

La mediana ya que podría mostrarnos un eje de equilibrio entre los datos. (Estudiante 4).

En general, las respuestas de los futuros profesores evidencian falta de comprensión sobre el tema, explicitando que la mediana permite resumir información cuantitativa, como en el caso de variables de naturaleza cualitativa ordinal.

En cambio, un 42,1% considera que la moda es el índice de resumen numérico adecuado, siendo justificada correctamente por un 15,8% (datos cualitativos). Algunos argumentos entregados son:

Para representar cual es la mejor nota, la mejor es la moda saber cuántas veces se repite mejor sirve (sobresaliente) y expresarlo en porcentaje en relación con la cantidad de alumnos de cada grupo, así podemos saber cuánto porcentaje de cada grupo obtuvo el mejor nivel (Estudiante 15).

La medida más apropiada sería la moda, el grupo 1 su moda es sobresaliente y el grupo 2 es bi-modal lo que muestra una muestra menos consistente (Estudiante 14).

Sería la mediana pues el promedio no se puede utilizar y la moda no representa de buena manera al curso, la mediana puede representar las dos mitades (Estudiante 9).

En resumen, los futuros profesores evaluados desconocen las propiedades teóricas fundamentales de los estadísticos asociados a los índices de resumen numérico. Esto se puede explicar porque normalmente se hace uso de la media aritmética, la mediana y la moda sin cuestionar la naturaleza de la variable a resumir numéricamente.

En coherencia con lo anterior se consulta a los encuestados *¿Cuál es el grupo que presenta mejores notas?* Esta pregunta puede ser respondida haciendo uso de algún estadístico que permita resumir información cualitativa, tal como la moda o la mediana (cualitativa ordinal). Además, la respuesta puede derivarse del análisis del resumen de información, al comparar la distribución de las notas mediante una tabla de distribución de frecuencia. En este caso, el 42,1% de los encuestados declaran que los integrantes del grupo 1 presentan mejores calificaciones., entregando argumentos con los siguientes:

El grupo que obtuvo mejores calificaciones es el grupo 1, ya que este presenta mayor número de aprobados y sobresalientes, la moda en el grupo 1 son los sobresalientes y en el grupo 2 estaría entre insuficiente y notables, lo que evidencia las mejores calificaciones del grupo 1 (Estudiante 2).

El grupo que obtuvo mejor nota fue el grupo 1, ya que este no se distribuye de forma normal (Estudiante 13).

Frente a estas respuestas, observamos que el estudiante 2 hace un análisis comparativo entre la distribución de las calificaciones de ambos grupos (Grupo 1 y Grupo 2) como así también de la moda y la mediana. Esto le permite tomar una decisión, logrando una respuesta de tipo acertada. En cambio, el estudiante 13 hace mención a la distribución normal, la cual modela información de tipo cuantitativa, aún cuando se está en presencia de datos cualitativos. Dicha situación podría ser explicada debido a experiencias previas a su formación docente. Por ejemplo, dentro de la asignatura de Estadística se incluye una unidad de estimación puntual de parámetros, en la que se trabaja con la media poblacional y se hace uso del método de máxima verosimilitud.

Esto significa, realizar estimaciones bajo el supuesto que se extrae una muestra de una población con distribución normal. Sin embargo, a nivel empírico, no siempre tiene probabilidad 1, lo cual se crea en los estudiantes la expectativa o el falso conocimiento de que siempre es posible tener datos que se puedan distribuir de manera normal. En cambio, un 15,8% declara que el grupo 2 obtiene mejores calificaciones. Entre los argumentos entregados se tienen:

El grupo 2 porque tienen menor varianza y además la mitad de los datos están sobre A. Mientras que el grupo 2 hay mayor cantidad de A (Estudiante 1).

El grupo 2 obtuvo mejores notas, puesto que cada grupo tiene diferente número de estudiantes y si vemos la proporción entre los dos, el grupo 2 es el que posee mejor calificación (Estudiante 5).

Al respecto, destaca el argumento dado por el estudiante 1, en que se hace mención a la varianza para dar respuesta a la interrogante, medida de variabilidad que teóricamente no es aplicable a variables de naturaleza cualitativa, como es el caso bajo estudio.

Índices de resumen numérico y sus propiedades

En Matemática, y en particular en Estadística, es de interés conocer las propiedades de los objetos estudiados. En este sentido, en la tercera actividad, se solicitó a los futuros profesores que declaren propiedades algebraicas, numéricas y estadísticas asociadas a los índices de resumen numérico introducidos a nivel de sistema escolar; a saber: media, mediana y moda.

Como se observa en la Tabla 3, el 26,3% de los futuros profesores presentes en el estudio declaran que el operador media aritmética no verifica la propiedad de ser una operación interna y afirman que posee la propiedad asociativa. En particular, llama la atención este resultado, pues tal como lo hace notar Mevarech (1983), la media aritmética presenta la estructura de grupo aditivo.

Tabla 3. Propiedades algebraicas, estadísticas y numéricas del objeto matemático media aritmética

Propiedad		Porcentaje (n=19)
Algebraica	No es una operación interna	26,3
	Conserva los cambios de origen y escala	5,3
	No presenta la propiedad Asociativa	5,3
	Presenta la propiedad Asociativa	5,3
	Presenta la propiedad Conmutativa	5,3
	No posee elemento neutro aditivo	5,3
	No posee elemento inverso aditivo	5,3
Estadística	La suma de las desviaciones de las observaciones respecto al promedio es cero.	10,5
	Es afectado por observaciones extremas	15,8
	En distribuciones simétricas coincide con la mediana y moda	5,3
	En distribuciones asimétricas es preferible utilizar la mediana o la moda	5,3
Númerica	El promedio se encuentra entre el mínimo y el máximo de los datos	15,8

Dentro de las propiedades estadísticas asociadas al promedio, el 10,5% de los estudiantes hacen referencia a la propiedad de que la *suma de las desviaciones de las observaciones respecto de la media es cero*; y un 15,8% que el promedio *es afectado por observaciones extremas*. Así también, un 15,8% explicitan que el promedio se encuentra entre el mínimo y máximo del conjunto de datos del cual ha sido calculado. En relación con este tipo de respuestas, es posible resumir que los estudiantes, en general, no hacen referencia a la naturaleza de la variable y de la unicidad del estadígrafo, situación no menor al momento de su cálculo.

Tabla 4. Propiedades algebraicas, estadísticas y numéricas del objeto matemático mediana

Propiedad		Porcentaje (n=19)
Algebraica	No es una operación interna	15,8
	Conserva los cambios de origen y escala	5,3
	No presenta la propiedad Asociativa	10,5
	No posee elemento neutro aditivo	5,3
	No posee elemento inverso aditivo	5,3
Estadística	No es afectada por observaciones extremas	21,1
	En distribuciones simétricas coincide la media, mediana y moda	5,3
	En distribuciones asimétricas es preferible utilizar la mediana o la moda	5,3
Numérica	La mediana se encuentra entre el mínimo y el máximo de los datos	15,8

En la Tabla 4 se muestran las respuestas emitidas por los profesores en formación con respecto a las propiedades de la mediana. Un 15,8% declara que este estadístico no es una operación interna; además los sujetos asumen que tiene la cualidad de no ser asociativa (10,5%), y no poseer elemento neutro o inverso.

De acuerdo con las propiedades estadísticas, el 21,1% de los encuestados expresan que este estadístico *no es afectado por observaciones extremas* (o datos atípicos). Sin embargo, no explicitan que en el caso que los datos se distribuyan de manera asimétrica es preferible considerarlo como valor representativo. Finalmente, con relación al estadístico moda (Tabla 5), se tiene que el 15,8% de los encuestados indica que es el valor que presenta la mayor frecuencia observada y, por otra parte, un 10,5% declara que no es única.

Tabla 5. Propiedades algebraicas, estadísticas y numéricas del objeto matemático moda

Propiedad		Porcentaje (n=19)
Algebraica	Es una operación interna	21,1
	No es una operación interna	10,5
	No presenta la propiedad Asociativa	5,3
	No posee elemento neutro aditivo	5,3
	No posee elemento inverso aditivo	5,3
	Conserva los cambios de origen y escala	5,3
Estadística	No es afectada por observaciones extremas	5,3
	En distribuciones simétricas coincide la media, mediana y moda	5,3
	En distribuciones asimétricas es preferible utilizar la mediana	5,3
	Es única	5,3
	No es única	10,5
Numérica	La moda se encuentra entre el mínimo y el máximo de los datos	10,5

Comentarios finales

A nivel país, el currículo escolar explicita contenidos y habilidades que deben desarrollar los estudiantes durante su escolarización. En tal sentido, la formación del profesorado debe dar respuesta a estas expectativas; recibiendo una instrucción pertinente y actualizada, cuya responsabilidad recae en la institución de Educación Superior que imparte esta carrera. Más aún cuando, las políticas de aseguramiento de la

calidad están el proceso de actualización de planes de estudio, la revisión de perfiles de ingreso y egreso, el alineamiento con modelos educativos centrados en el estudiantado, entre otros.

En particular, la carrera de Pedagogía en Educación Matemática de la cual provienen los sujetos en estudio se encuentra acreditada y en proceso de renovación curricular. En este contexto, estudios empíricos resultan de interés porque entregan resultados para la toma de decisiones sobre el diseño del futuro itinerario formativo. Es así como, a la luz de los resultados del presente estudio, un área de interés resulta el ámbito de Estadística, en el cual los estudiantes evaluados han demostrado un bajo nivel de dominio conceptual y teórico acerca de los índices de resumen numérico. Esto último, pese a considerar un número reducido de futuros profesores, resulta preocupante porque los estudiantes se encuentran en el ciclo final de su formación y contarían de manera insuficiente con las habilidades necesarias para dar respuesta a los requerimientos demandados a nivel país en el eje Datos y Probabilidad; pues es sabido que no se puede enseñar lo que no se domina (Hernández y Sancho, 1996).

En segundo lugar, a nivel de estadística descriptiva, uno de los contenidos que son introducidos en etapas tempranas en el sistema escolar son los índices de resumen numérico (media aritmética, mediana y moda), los que se asocian a un *valor típico*, es decir, valor central más pertinente para describir un conjunto de datos (Russel y Mokros, 1991). Esto implica desafíos para la enseñanza, pues el futuro profesor debe ser capaz de transmitir la idoneidad del índice en relación a la naturaleza de las variables y el contexto; asumiendo que estas medidas juegan un rol fundamental en el resumen. Al respecto, las respuestas de los futuros profesores bajo estudio desde la alteridad, permiten afirmar que estos tienden a replicar una mirada algorítmica en relación al uso de la media aritmética, la mediana y la moda; omitiendo consideraciones teóricas que permitan fundamentar la toma de decisiones en determinados contextos. Dicha situación es alarmante, dado que estos futuros profesores son los encargados de implementar cambios curriculares en el sistema escolar, en forma alineada con las directrices curriculares, centrado en potenciar el razonamiento estadístico. El presente estudio, en este sentido, deja abierta la posibilidad de analizar las respuestas de un grupo más amplio de futuros profesores, asunto que constituye una de las limitaciones de investigaciones del tipo transversal.

Dada la actual estructura curricular chilena, en octavo año de Educación Primaria, se introducen los diagramas de cajas, representación que permite visualizar características de datos cuantitativos tales como asimetría, tendencia central, variabilidad y valores atípicos; estos últimos se presentan de manera frecuente en datos reales. En este contexto, el tratamiento de dichos valores exige tener conocimientos conceptuales de estadística debido a que ello incide en la toma de decisiones al momento de hacer uso de algún estadígrafo de resumen. A nivel de resultados se evidencia que el grupo en estudio, de futuros profesores de Educación Secundaria, presenta porcentajes de logro descendidos al respecto, así como argumentaciones teóricamente incorrectas, situación que devela la falta de actividades relacionadas con datos empíricos y su tratamiento estadístico.

Finalmente, los hallazgos del presente estudio entregan evidencia, que de ser contrastado con un grupo mayor, permitiría reorientar los programas de asignatura, considerando la importancia de una sólida mirada conceptual del eje Estadística y Probabilidad que permita tomar decisiones acertadas al momento de hacer uso de los índices de resumen numérico, más allá del mero uso algorítmico.

Agradecimientos

Investigación Financiada por Dirección de Investigación Universidad del Bío-Bío (DIUBB) código 166523 3/R: “Habilidades de análisis e interpretación de datos en presencia de variabilidad estadística: un estudio comparativo en estudiantes de formación inicial docente y profesores del sistema escolar”.

Referencias

- Arteaga, P., Batanero, C., Contreras, J.M. y Cañadas, G. (2016). Evaluación de errores en la construcción de gráficos estadísticos elementales por futuros profesores. *Revista Latinoamericana de Investigación en Matemática Educativa*, 19(1), 15-40.
- Ávalos, B. (2014). La formación inicial docente en Chile: tensiones entre políticas de apoyo y control. *Estudios Pedagógicos*, 40(Especial), 11-28.
- Barr, G. V. (1980). Some student ideas on the median and the mode. *Teaching Statistics*, 2(2), 38-41.
- Cobo, B. y Batanero, C. (2000). La mediana en la educación secundaria obligatoria: ¿Un concepto sencillo? *UNO. Revista de Didáctica de las Matemáticas*, 23, 85-96.
- Díaz-Levicoy, D., Sepúlveda, A., Vásquez, C. y Opazo, M. (2016). Lectura de tablas estadísticas por futuras maestras de Educación Infantil. *Educação Matemática Pesquisa*, 18(3), 1099-1115.
- Estrada, A., Batanero, C. y Fortuny, J. (2004). Un estudio sobre conocimientos de estadística elemental de profesores en formación. *Educación Matemática*, 16(1), 89-111.
- Estrella, S. (2016). Comprensión de la media por profesores de educación primaria en formación continua. *REDIE. Revista Electrónica de Investigación Educativa*, 18(1), 1-22.
- Estrella, S., Olfos, R. y Mena-Lorca, A. (2015). Pedagogical knowledge of statistics content among primary school teachers. *Educação e Pesquisa*, 41(2), 477-493.
- Gal, I. (2002). Adult's statistical literacy: Meaning, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Garfield, J. y Konold, C. (1992). *Statistical reasoning assessment, Part 2: Statistics in context*. Minnesota: National Science Foundation.
- Garrett, A. y García, J.A. (2008). Caracterización de la comprensión de algunos aspectos de la media aritmética: un estudio con alumnos de secundaria y universitarios. *Enseñanza de la Matemática*, 17(1), 31-57.
- Gea, M.M., Arteaga, P. y Cañadas, G.R. (2017). Interpretación de gráficos estadísticos por futuros profesores de Educación Secundaria. *Avances de Investigación en Educación Matemática*, 12, 19-37.

- Garfield, J. y Konold, C. (1992). *Statistical reasoning assessment. Part 2: Statistics in context*. Minnesota, MN: National Science Foundation.
- Godino, J.D. (1999, Septiembre). *Análisis epistémico, semiótico y didáctico de procesos de instrucción matemática*. Trabajo presentado en el grupo de trabajo “La didáctica de la matemática como disciplina científica” en el III Simposio de la Sociedad Española de Investigación en Educación Matemática, Valladolid, España.
- Hacking, I. (2005). *El surgimiento de la probabilidad*. Barcelona, España: Gedisa.
- Hernández, F. y Sancho, J. (1996). *Para enseñar no basta con saber la asignatura*. Barcelona, España: Paidós.
- Ho, F., Kee, G. y Ramakrishnan, C. (2017). *Texto del estudiante Matemática 5° Básico*. Santiago, Chile: Santillana.
- Jacobbe, T. (2012). Elementary school teachers’ understanding of the mean and median. *International Journal of Science and Mathematics Education*, 10(5), 1143-1161.
- Jacobbe, T. y Carvalho, C. (2011). Teachers’ understanding of averages. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 199-209). Netherlands: Springer.
- Konold, C. y Pollatsek, A. (2004). Conceptualizing an average. En D. Ben-zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 169-199). Netherlands: Springer.
- Marín, F., Castillo, M. y Fernández, P. (2011). *Texto de Estudiante Matemática 5° Básico*. Santiago, Chile: Santillana.
- McMillan, J. H. y Schumacher, S. (2005). *Investigación educativa*. Madrid: Pearson-Adisson Wesley.
- Mevarech, Z. (1983). Deep structure model of students’ statistical misconceptions. *Educational Studies in Mathematic*14(1), 415-429.
- Merino, R., Muñoz, V., Pérez, B. y Rupin, P. (2016). *Texto del estudiante Matemática 7° básico*. Santiago, Chile: SM.
- MINEDUC. (2009). *Objetivos fundamentales y contenidos mínimos obligatorios de la Educación Básica y Media*. Santiago, Chile: Gobierno de Chile.
- MINEDUC. (2012a). *Estándares orientadores para carreras de pedagogía en educación media. Estándares pedagógicos y disciplinarios*. Santiago, Chile: Autor.
- MINEDUC. (2012b). *Matemática educación básica. Bases curriculares*. Santiago, Chile: Unidad de Currículum y Evaluación.
- MINEDUC. (2016). *Decreto Ley n° 20.903, de 2016. Sistema de Desarrollo Profesional Docente y Modifica Otras Normas*. Valparaíso: Congreso Nacional.
- Moore, D. (1991). Teaching statistics as a respectable subject. En F. Gordon y S. Gordon (Eds.), *Statistics for the twenty-first century* (pp. 14-25). Washington: Mathematical Association of America.
- Ortiz, J. J., Font, V. y Mayén, S. (2009). Significados personales de la media aritmética de profesores en formación. En M. J. González, M. T. González y J. Murillo (Eds.), *Investigación en Educación Matemática XIII* (pp. 345-353). Santander: SEIEM.

- Rodríguez-Alveal, F. (2017). Alfabetización estadística en profesores de distintos niveles formativos. *Educação & Realidade*, 42(4), 1459-1477.
- Rodríguez-Alveal, F., Maldonado-Fuentes, A. y Sandoval, P. (2016). Comprensión de las medidas de tendencia central: un estudio comparativo en estudiantes de pedagogía en matemática en dos instituciones formadoras chilenas. *Avaliação: Revista da Avaliação da Educação Superior*, 21(3), 929-952.
- Rodríguez-Alveal, F. y Sandoval, P. (2012). Habilidades de codificación y descodificación de tablas y gráficos estadísticos: un estudio comparativo en profesores y alumnos de pedagogía en enseñanza básica. *Avaliação: Revista da Avaliação da Educação Superior*, 17(1), 207-235.
- Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1010). Greenwich: Information Age y NCTM.
- Sotomayor-Echenique, C., Coloma-Tirapegui, C.J., Parodi-Sweis, G., Ibáñez-Orellana, R., Cavada-Hrepich, P. y Gysling-Caselli, J. (2013). Percepción de los estudiantes de pedagogía sobre su formación inicial. *Magis. Revista Internacional de Investigación en Educación*, 5(11), 375-392.
- Vaillant, D. (2007). Mejorando la formación y el desarrollo profesional docente en Latinoamérica. *Pensamiento Educativo. Revista de Investigación Educativa Latinoamericana*, 41(2), 207-222.
- Vezub, L.F. (2016). Los saberes docentes en la formación inicial. La perspectiva de los formadores. *Pensamiento Educativo. Revista de Investigación Educativa Latinoamericana*, 53(1), 1-14.
- Watson, J. (1997). Assessing statistical literacy through the use of media surveys. En I. Gal y J.B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107-121). Amsterdam: IOS Press.