

Algunas nociones prioritarias en el estudio del muestreo

Some priority notions in the study of sampling

Amable Moreno

Universidad de Mendoza, Argentina

Resumen

En este trabajo analizamos algunas ideas prioritarias en la comprensión del muestreo en estudiantes universitarios. Nos referimos a las ideas de muestra aleatoria simple, estadístico y a la distribución de la media muestral. La aplicación de un cuestionario con diez ítems a un grupo de treinta y cuatro estudiantes universitarios nos permitió detectar algunas dificultades en la interpretación de estas nociones. Las dificultades más importantes se refieren a la falta de interpretación de las nociones de muestra aleatoria, la media de la media muestral y el efecto del tamaño de la muestra sobre la variabilidad de la media muestral.

Palabras clave: población, muestra, estadístico, distribución

Abstract

In this paper we analyze some priority ideas in the understanding of sampling in university students. We refer to the ideas of simple random sample, statistical and the distribution of the sample mean. The application of a questionnaire with ten items to a group of thirty-four university students, allowed us to detect some difficulties in the interpretation of these notions. The most important difficulties relate to the lack of interpretation of the notions of random sample, the mean of the sample mean and the effect of the sample size on the variability of the sample mean.

Keywords: population, sample, statistical, distribution

1. Introducción

Para algunos autores uno de los problemas didácticos en la enseñanza de la estadística en la universidad es la enseñanza de la inferencia (Artigue, Batanero y Kent, 2007; Moore, 1997).

El muestreo es un tema fundamental en la comprensión posterior de la inferencia estadística pero, sin embargo, la investigación previa describe numerosos errores en la comprensión por parte de los estudiantes (Begué, Batanero, Ruiz y Gea, en prensa). En parte esta dificultad se debe a la necesidad de manejar simultáneamente tres distribuciones (Harradine, Batanero y Rossman, 2011; Kadijevich, Kokol-Voljc y Lavicza, 2008) que están presentes durante todo el proceso; la distribución poblacional, la distribución de una muestra y la distribución del estadístico en el muestreo. Además, de acuerdo con Batanero, Díaz, Contreras y Roa (2013) y Harradine et al. (2011), el razonamiento inferencial requiere de la coordinación de distintos conceptos: los datos, la población de donde se tomaron los datos y las posibles muestras de la misma.

Este trabajo se inicia a partir de la necesidad de conocer las ideas de los estudiantes sobre distribuciones muestrales, antes de comenzar con la enseñanza de la inferencia estadística. Las ideas relativas a muestras y distribuciones muestrales inciden en el estudio de la inferencia estadística. En efecto, desde nuestra experiencia docente hemos podido detectar que los estudiantes realizan correctamente los procedimientos para la construcción de intervalos de confianza, y la verificación de hipótesis estadísticas; sin embargo, tienen escasa comprensión conceptual del proceso subyacente. Por este

motivo realizamos el presente estudio, cuyos resultados parciales presentamos en esta comunicación.

El estudio se realizó con treinta y cuatro estudiantes universitarios de segundo año, de ciencias económicas. Consiste en la aplicación de un cuestionario con diez ítems, referidos a las ideas muestra aleatoria, estadístico y distribución muestral de la media.

2. Marco teórico

Siguiendo a Harradine et al. (2011), la inferencia está constituida por tres elementos fundamentales: proceso de razonamiento, los conceptos y los cálculos; mientras que los cálculos se ven facilitados por los ordenadores, la comprensión de los conceptos, entre los que mencionan la aleatoriedad y el muestreo no es completa.

El tema de la distribución muestral ha sido investigado por varios autores. Saldanha y Thompson (2002) realizaron un estudio sobre las ideas de muestra, distribución muestral y márgenes de error, en un grupo de estudiantes de 16 a 18 años. La investigación se centró en el análisis de tres niveles del concepto de distribución (distribución de la población, distribución de la muestra y distribución muestral). Los estudiantes no logran diferenciar los tres niveles del concepto de distribución; y además los autores encontraron que la mayoría de los estudiantes tienen una imagen aditiva de la muestra, porque tienden a ver a la muestra aleatoria como un subconjunto de la población. Desde otro punto de vista, Moreno (2018a, 2018b) recomienda analizar cuatro niveles en el concepto de distribución: distribución de la poblacional, distribución de la muestra observada, distribución de la muestra aleatoria como variable n-dimensional y distribución muestral de un estadístico.

Por otra parte, Chance, delMas y Garfield (2004) sugieren reflexionar sobre los conceptos sobre los cuales se basan los procedimientos inferenciales (muestra, estadístico, variabilidad de las muestras y el efecto del tamaño sobre la variabilidad). Estos autores realizaron un estudio con estudiantes que iban a iniciar estudios superiores, a lo largo de un curso en el que usan un software. Los resultados muestran que los estudiantes no logran explicar los conceptos que están en juego en los procedimientos que aplican. En particular, no comprenden los conceptos de distribución muestral y variabilidad muestral

Vanhoof, Castro Sotos, Onghena y Verschaffel (2007) realizaron una investigación sobre la distribución muestral con un grupo de estudiantes universitarios que cursaban el segundo año de la carrera, en la que analizaron la efectividad del uso de un software denominado Sampling Activity Distribution. Los resultados revelaron una mejora significativa en la comprensión de la distribución muestral, e indicaron que solamente la mitad de los estudiantes mostraron una comprensión adecuada de la distribución muestral.

Wild, Pfannkuch, Regan y Horton (2011) y otros autores han señalado que, de la colección de artículos encontrados en dicho momento en la literatura de la investigación, se puede concluir que, para ayudar a los estudiantes a desarrollar ideas sobre muestreo, se requiere mucho más tiempo que el asignado en el curriculum de las diferentes carreras. Pensamos que es importante introducir las ideas de muestra y muestreo a estudiantes tempranamente en su aprendizaje estadístico, para que, llegado el momento, en el nivel superior estén preparados para el estudio de las ideas formales de muestra aleatoria y distribuciones muestrales. Además, los estudiantes deberían tener

una buena comprensión de los conceptos fundamentales de muestra, variabilidad, distribución y centro.

Por otra parte, para Chance, DelMas y Garfield (2004, pp. 301), los estudiantes de los cursos introductorios de estadística del nivel universitario, deben comprender acerca de las distribuciones muestrales :

- Una distribución muestral de medias muestrales es una distribución de todas las posibles medias muestrales de muestras del mismo tamaño seleccionadas de una población con media μ y desviación estándar σ ; ésta es una distribución de probabilidad para la media muestral.
- La distribución muestral de la media tiene la misma media que la población.
- A medida que el tamaño de la muestra aumenta la variabilidad muestral se reduce.
- El error estándar de la media es una medida de variabilidad de los valores de las medias muestrales.
- Diferentes tamaños muestrales conducen a diferentes probabilidades para el mismo valor del estadístico.
- Las distribuciones muestrales tienden a tener la forma de una distribución normal en lugar de la forma de la distribución de la población, incluso para muestras pequeñas.
- A medida que los tamaños muestrales se vuelven más grandes, todas las distribuciones muestrales para la media se parecen (es decir, tienen la misma forma) independientemente de la población de la cual se extraen.
- Las medias muestrales son más normales y menos variables que las observaciones individuales.
- Ser capaz de distinguir entre una distribución de observaciones en una muestra y una distribución de medias muestrales de muestras de tamaño n ($n > 1$) que han sido seleccionadas al azar.

Finalmente, citamos el trabajo de Begué, Batanero y Gea (2018), quienes piden a un grupo de estudiantes de educación secundaria (14 y 16 años de edad) formar muestras de distribuciones binomiales, variando el valor de la proporción muestral y del tamaño de la muestra. Los autores concluyen que los estudiantes comprenden la idea de valor esperado en la distribución muestral, pero no son conscientes de la relación entre el tamaño de la muestra y la variabilidad de la distribución muestral. De hecho, conceden mayor variabilidad a las muestras de mayor tamaño, en contra de lo que indica la teoría estadística.

3. Metodología

En este trabajo se realizó un estudio exploratorio sobre la comprensión del concepto de muestra aleatoria simple, estadístico y nociones asociadas a la distribución muestral de la media por parte de un grupo de 34 estudiantes universitarios de segundo año de la carrera de ciencias económicas, que se encuentran realizando por primera vez un curso introductorio de Estadística, sin conocimientos previos de la asignatura. Para la realización del mismo, se recurrió a un cuestionario compuesto por diez ítems, que se describe a continuación, y que fue completado luego de que los estudiantes recibieran

instrucción sobre muestreo. En cada ítem incluido a continuación el estudiante debía indicar si la proposición es verdadera o falsa.

Ítem 1: A toda muestra aleatoria observada le corresponde una distribución de frecuencias relativas.

Ítem 2: Las variables aleatorias (X_1, X_2, \dots, X_n) constituyen una muestra aleatoria de tamaño n , si se cumple:

- a) Las X_i son variables aleatorias independientes;
- b) todas las X_i tienen la misma distribución de probabilidad, y esta distribución de probabilidad es la misma que tiene la población de la cual se extrajo

Ítem 3: "En toda muestra aleatoria simple (X_1, X_2, \dots, X_n) se verifica:

- a) $F(X_1)=F(X_2)=\dots=F(X_n)$
- b) $F(X_1, X_2, \dots, X_n) = F(X_1).F(X_2).\dots.F(X_n)$

Ítem 4: En una muestra aleatoria observada las medidas descriptivas: media, mediana, varianza y rango; coinciden exactamente con las correspondientes de la población.

Ítem 5: Un estadístico es una función de las variables aleatorias de la muestra, por lo tanto es una variable aleatoria.

Ítem 6: Dada una variable aleatoria X , su imagen coincide con la imagen de la media muestral, para muestras de tamaño $n > 1$.

Ítem 7: La probabilidad de que la media muestral tome valores que se alejan de la media poblacional en menos de un desvío poblacional, es mayor en muestras de tamaño 25 que en muestras de tamaño 100.

Ítem 8: El tamaño muestral debe ser grande para que el valor de la media muestral se ubique en la cola de la distribución poblacional.

Ítem 9: El error estándar de la media muestral nunca toma valores menores o iguales a la desviación estándar poblacional.

Ítem 10: La media de una variable que tiene distribución asimétrica positiva será más grande que la media de la distribución muestral de la media para muestras grandes de tamaño n , tomadas de la población considerada.

En el ítem 1 el estudiante debe reconocer que a toda muestra aleatoria observada le corresponde una distribución de frecuencias relativas, que indica cómo se distribuyen los valores de la variable en la muestra observada.

El objetivo del ítem 2 se refiere al reconocimiento de la definición de muestra aleatoria simple, como una variable aleatoria n -dimensional, que tiene una función densidad de probabilidad conjunta, y que cumple dos propiedades.

En el ítem 3 se pretende nuevamente que reconozca la definición de muestra aleatoria, como una variable aleatoria n -dimensional, pero en este caso las dos propiedades se expresan en términos de la función de distribución acumulada conjunta.

En el ítem 4 los estudiantes deben diferenciar algunos estadísticos muestrales de los correspondientes parámetros poblacionales.

El ítem 5 plantea la definición de estadístico, y el objetivo es que los estudiantes reconozcan la naturaleza aleatoria de un estadístico

El ítem 6 se refiere al conjunto de valores que toma la media muestral y lo compara con el conjunto de valores que toma la variable poblacional asociada, los estudiantes deberían reconocer que la imagen de la variable X no coincide con la imagen de la variable \bar{X} .

El ítem 7 se pretende conocer en qué medida los estudiantes reconocen la variabilidad de la media muestral; es decir, reconocer cómo influye el tamaño muestral en la distribución de la media muestral. Por lo tanto, diferentes tamaños muestrales conducen a diferentes distribuciones de probabilidad para el estadístico media muestral.

El ítem 8 tiene el mismo objetivo que el ítem anterior, reconocer la influencia del tamaño muestral en la distribución de la media muestral.

El ítem 9 se refiere a la variabilidad de la media muestral, al igual que los dos ítems anteriores.

El ítem 10 implica el reconocimiento de la propiedad de la media muestral como estimador insesgado de la media poblacional.

4. Resultados

La Tabla 1 muestra la cantidad y porcentaje de respuestas correctas, respuestas incorrectas y no contestadas para cada uno de los ítems.

A continuación, presentamos cada uno de los ítems con los resultados obtenidos en cada caso.

Tabla 1. Frecuencia (y porcentaje) de respuestas a cada pregunta

Ítem	Correcta	Incorrecta	No Contesta
Ítem 1	24 (70)	9 (27)	1 (3)
Ítem 2	18 (53)	14 (41)	2 (6)
Ítem 3	5 (15)	25 (73)	4 (12)
Ítem 4	31 (91)	3 (9)	0 (0)
Ítem 5	28 (80)	6 (17)	0 (0)
Ítem 6	21 (60)	10 (29)	3 (9)
Ítem 7	12 (35)	19 (56)	3 (9)
Ítem 8	18 (53)	14 (41)	2 (6)
Ítem 9	20 (59)	10 (29)	4 (12)
Ítem 10	16(47)	17(50)	1 (3)

En el ítem 1 el estudiante debe reconocer que a toda muestra aleatoria observada le corresponde una distribución de frecuencias relativas, que indica cómo se distribuyen los valores de la variable en la muestra observada. Esta idea es aceptada como correcta por 24 (70%) estudiantes. Mientras que, 9 (27%) estudiantes no reconocen la existencia de una distribución; y solamente 1(3%) de los estudiantes no contestan. La mayoría de los estudiantes reconocen la existencia de una distribución de frecuencias relativas para cada muestra aleatoria observada; y unos pocos no reconocen dicha distribución porque ven a la muestra como un conjunto de datos aislados y no como una estructura que nos da una idea acerca de la distribución de la población de la cual se extrajo.

En el ítem 2 se presenta la definición de muestra aleatoria simple, como una variable aleatoria n -dimensional, que tiene una función densidad de probabilidad conjunta, y que cumple dos propiedades. Esta definición es aceptada como verdadera por 18 (53%) estudiantes; mientras que 14 (41%) no reconoce la definición de muestra aleatoria, como una variable aleatoria conjunta, que cumple dos propiedades; y solamente 2 (6) estudiantes no contestan a este ítem.

En el ítem 3, se presenta nuevamente la definición de muestra aleatoria, como una variable aleatoria n -dimensional, pero las dos propiedades se expresan en términos de la función de distribución acumulada conjunta. Encontrando un bajo porcentaje de respuestas correctas; solamente 5 (15%) estudiantes contestan correctamente, mientras que 25 (73%) estudiantes no reconocen la definición de muestra aleatoria cuando se da en términos de la función de distribución acumulada conjunta; y 4 (12%) no contestan. Es evidente que los estudiantes no asocian la función densidad de probabilidad conjunta con la función de distribución acumulada conjunta; y no alcanzan a comprender que la independencia estadística se puede definir en términos de la función densidad de probabilidad conjunta o en términos de la función de distribución acumulada.

Con el ítem 4 se pretende que los estudiantes diferencien algunos estadísticos muestrales de los parámetros poblacionales. Los resultados muestran que, en general los estudiantes logran diferenciar estadísticos de parámetros, dado que 31 (91%) estudiantes lo hacen de manera correcta, 3 (9%) estudiantes no diferencian los estadísticos de los parámetros y solamente 1 (3%) estudiante no contesta.

En el ítem 5, los estudiantes reconocen la naturaleza aleatoria de un estadístico; y por lo tanto un alto porcentaje, el 28 (80%) estudiantes contestan correctamente, 6 (17%) contestan de manera incorrecta y ningún estudiante no contesta.

Más de la mitad de los estudiantes 21 (60%) reconoce en el ítem 6 la diferencia entre el conjunto de valores que toma una variable X y el conjunto de valores que puede tomar una media muestral \bar{X} para muestras de igual tamaño n , extraídas de la misma población. Mientras que, solamente 10 (29%) estudiantes no logran reconocer esta diferencia; y 3 (9%) no contestan.

Con el ítem 7 se pretende conocer, en qué medida los estudiantes reconocen la variabilidad de la media muestral; es decir, reconocer cómo influye el tamaño muestral en la distribución de la media muestral. Por lo tanto, diferentes tamaños muestrales conducen a diferentes distribuciones de probabilidad de la media muestral. Los resultados muestran que solamente 12 (35%) de los estudiantes interpretan correctamente la variabilidad de la media muestral; mientras que 19 (56%) estudiantes no reconocen la influencia del tamaño de la muestra en la distribución de la media muestral; y 3 (9%) estudiantes no contestan.

El ítem 8 tiene el mismo objetivo que el ítem anterior, sin embargo el número de respuestas correctas es algo mayor; dado que 18 (53%) estudiantes responde correctamente 14 (41%) estudiantes no reconoce que la media muestral es una variable aleatoria con una variabilidad que depende del tamaño de la muestra y 2 (6%) estudiantes no contesta. Sin embargo, el nivel de complejidad de este ítem es menor que el del ítem anterior.

El ítem 9 tiene el mismo objetivo que los dos ítems anteriores, sin embargo el número de respuestas correctas es mayor que en los dos ítems anteriores, siendo de 20 (59%) estudiantes los que reconocen la variabilidad de la media muestral; 10 (29%) los que no

interpretan correctamente la variabilidad de la media muestral y 4 (12%) estudiantes no contestan. La dificultad de este ítem es menor que la dificultad de los dos ítems anteriores.

En el ítem 10, los estudiantes deben reconocer que la media muestral tiene una distribución de probabilidad, y que la media de todas las medias muestrales es la media poblacional. Un poco menos de la mitad de los estudiantes reconocen esta propiedad de la media muestral, que es fundamental en la comprensión de la distribución muestral de la media, dado que 16 (47%) estudiantes responden correctamente, 17 (50%) no logran reconocer que la media muestral tiene una distribución de probabilidad y que la media de todas las medias muestrales es la media poblacional; y sólo 1(3%) no contestan.

Los ítems 7, 8 y 9 tienen el mismo objetivo, sin embargo el porcentaje de respuestas correctas es diferente. Por lo que podemos concluir que los respuestas dependen de cómo se planteen los enunciados. A medida que aumenta la complejidad de los enunciados, aumenta el porcentaje de respuestas correctas. De esta manera, podemos conocer el nivel de comprensión que alcanzan los estudiantes; por lo que se concluye que solamente 12 (35%) de los estudiantes logra una comprensión adecuado de la influencia del tamaño muestral en la distribución de probabilidad de la media muestral.

El ítem 3 es el que presenta mayor dificultad debido a que los estudiantes no logran interpretar correctamente el concepto de independencia estadística y aplicarlo a la definición de muestra aleatoria.

En orden de dificultad al ítem 3 sigue al ítem 7, en el que también la dificultad aparece debido a la falta de integración de conocimientos. Dado que tamaños muestrales distintos dan lugar a distribuciones de probabilidad distintas para la media muestral; y por lo tanto a menor variabilidad de la media muestral le corresponde mayor probabilidad de la media muestral para valores fijos de \bar{X} .

Siguiendo el orden de dificultad se presenta el ítem 10, en el que para más de la mitad de los estudiantes, la asimetría de la distribución poblacional influye en la determinación de la media de la media muestral, desconociendo la propiedad de estimador insesgado de la media muestral.

5. Conclusiones y discusión

El cuestionario se aplicó a treinta cuatro estudiantes universitarios, luego de haber recibido instrucción sobre muestreo. Sin embargo, los resultados muestran discrepancias en relación con instrucción recibida sobre estos temas. Esto se debe en parte al poco tiempo establecido en el curriculum para la enseñanza de estos temas, y a la falta de disponibilidad de recursos tecnológicos apropiados para la enseñanza. Consideramos que se deberían aplicar otras estrategias didácticas para solventar estas discrepancias y destinar más tiempo a la enseñanza de estos temas.

Sabemos que para comprender la esencia de la inferencia estadística es necesario comprender la naturaleza de una población, una muestra aleatoria y de un estadístico; y que cada uno de ellos le corresponde una distribución de probabilidad, y estas distribuciones difieren entre sí. Sin embargo, los resultados de este estudio nos informan que, un número importante de estudiantes no reconocen la existencia de la distribución de los valores de una variable en la muestra observada. De lo que podemos inferir que, no está presente la idea de distribución, dado que no asocian la muestra observada a una distribución de frecuencias relativas, y que además puede ser caracterizada por un

conjunto de medidas que surgen del cálculo de unos estadísticos descriptivos; que permiten dotarla de una estructura compleja, y no concebir a la muestra observada como unos datos aislados.

Cuando se presenta la definición de muestra aleatoria, tal como se presentó en el curso, el 53% de los estudiantes reconocen la definición; pero cuando se presenta la definición de muestra aleatoria en términos de la función de distribución acumulada conjunta, el reconocimiento de la definición de muestra aleatoria baja en forma considerable; dado que solamente el 15% de estudiantes la reconocen correctamente. Esta discrepancia entre los resultados de los dos ítems, sugieren la idea de que unos pocos estudiantes logran integrar varios conceptos estadísticos: variable aleatoria conjunta, función densidad conjunta, distribución de distribución acumulada conjunta e independencia estadística.

En relación con el concepto de estadístico y parámetro, los estudiantes no tienen dificultad en reconocerlos y, por tanto, en diferenciarlos. Pero cuando se les solicita que comparen el conjunto de valores que toma una variable con el conjunto de valores que toma la media muestral para muestras de igual tamaño, el 60 % de los estudiantes logra reconocer la diferencia entre ambos conjuntos. Por lo tanto, podríamos concluir que un porcentaje importante de estudiantes no logra comprender la verdadera naturaleza del estadístico media muestral. En particular, cuando se considera la distribución muestral de la media, menos de la mitad de los estudiantes, 16 (47%) reconoce que la media de todas las medias muestrales es igual a la media de la población.

Finalmente, los ítems 6, 7 y 8 con distintos niveles de complejidad, a los cuales subyace la misma idea, que es la variabilidad de la media muestral; ha dado lugar a un número de respuestas correctas diferentes. Por lo que, podemos concluir que los resultados dependen, en muchos casos, de cómo se planteen los enunciados que se pretenden analizar.

Consideramos que estas dificultades que muestran los estudiantes se deben en parte, a lo que expresa Saldanha y Thompson (2002), si bien una muestra es una parte crítica de la inferencia estadística, "el muestreo no se ha caracterizado en la literatura como un esquema de ideas interrelacionadas que conlleva repetidas selecciones aleatorias, variabilidad y distribución" (pp. 258). En el mismo sentido, para Castro-Sotos (2009) los estudiantes deben desarrollar una comprensión de muchos conceptos probabilísticos y estadísticos y las relaciones entre ellos, antes de introducir a los estudiantes en la Inferencia Estadística. Es necesario que las ideas subyacentes se desarrollen a lo largo de años y no en unas semanas.

Así también, para Chance, DelMas y Garfield (2004) el razonamiento sobre la distribución de muestreo requiere que los estudiantes integren varios conceptos estadísticos y puedan razonar sobre el comportamiento de todas las muestras posibles, lo que implica un proceso de pensamiento que no es alcanzado por muchos estudiantes. De acuerdo con estos autores, muchos estudiantes no logran desarrollar una comprensión profunda del concepto de distribución de muestreo y sólo pueden adquirir un conocimiento mecánico de la inferencia estadística.

Para Batanero et al.(1994), los estudiantes deben entender que el propósito de extraer una única muestra requiere de la asimilación de dos ideas: representatividad muestral y variabilidad muestral. En estas situaciones, el propósito de extraer una muestra es cuantificar el nivel de inusual de esa muestra en relación con las muchas otras muestras que podrían haberse extraído.

Como se expresa en Harradine et al.(2011), se requiere de la creación y evaluación de un curriculum para el nivel de secundaria que desarrolle sistemáticamente las ideas clave que sustentan la inferencia estadística, y que proporcionen a los estudiantes una base adecuada para la instrucción formal. Estos autores sugieren introducir a los estudiantes en un curso de estadística informal, cuya característica sea que las actividades dadas involucren a los alumnos en el proceso de razonamiento inferencial sin depender de distribuciones de probabilidad y fórmulas. Para algunos autores (Rubin, Hammerman, y Konold, 2006) la inferencia informal es una colección de ideas que sustentan la inferencia estadística clásica. Entre estas ideas se incluyen, comprender las fuentes de variabilidad, reconocer el efecto del tamaño de la muestra y ser capaz de identificar tendencias y fuentes de sesgo. Para otros autores (Zieffler, Garfield, Delmas, y Reading; 2008) la inferencia informal incluyen ciertos tipos de razonamiento: razonamiento sobre las posibles características a partir de una muestra observada; razonamiento sobre las posibles diferencias entre dos poblaciones a partir de las diferencias entre dos muestras observadas; y si es probable o poco probable que un estadístico muestral proporcione una buena estimación de un parámetro poblacional.

Si bien, hay numerosos recursos informáticos que existen para hacer de estos conceptos difíciles más concretos(Biehler, Ben-Zvi, Bakker y Maker, 2013), y a pesar que los diseños curriculares presentan estos tópicos en el nivel de secundaria; no hay suficiente investigación que nos informe cómo llevar a cabo la instrucción estadística para el desarrollo del pensamiento estadístico. Tampoco se conoce mucho sobre la comprensión de los profesores sobre estos temas (Moreno, 2018a).

Como profesores debemos tener presente que el análisis de los datos estadísticos no es un proceso mecánico, y por lo tanto no debe ser enseñado ni aplicado de esta manera, ya que las estadísticas no son una forma de hacer, sino una forma de pensar (Batanero, 2000).

Referencias

- Artigue, M., Batanero, C. y Kent, P. (2007). Mathematics thinking and learning at post-secondary level. En F. Lester K. (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 1011-1049). Charlotte, NC: Information Age.
- Batanero, C. (2000). Controversies around significance tests. *Mathematical Thinking and Learning*, 2(1-2), 75–98.
- Batanero, C., Díaz, C., Contreras, J. M. y Roa, R. (2013). El sentido estadístico y su desarrollo. *Números*, 83, 7-18.
- Batanero, C., Godino, J. D., Vallecillos, A., Green, D. R. y Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.
- Begué, N., Batanero, C. y Gea, M. M. (2018). Comprensión del valor esperado y variabilidad de la proporción muestral por estudiantes de educación secundaria obligatoria. *Enseñanza de las Ciencias*, 36(2), 63-79.
- Begué, N., Batanero, C., Ruiz, K. y Gea, M.M. (En prensa). Investigaciones sobre comprensión del muestreo. *Boletín de Estadística e Investigación Operativa*.
- Biehler, R., Ben-Zvi, D., Bakker, A. y Maker, K. (2013). Technology for enhancing statistical reasoning at the school level. En M.A. Clements, A. Bishop, C. Keitel, J.

- Kilpatrick y F. Leung (Eds.), *Third international handbook of mathematics education* (pp. 643-690). New York: Springer.
- Chance, B., delMas, R. C. y Garfield, J. (2004). Reasoning about sampling distributions. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295–323). Amsterdam: Kluwer.
- Castro-Sotos, A. E. (2009). How confident are students in their misconceptions about hypothesis tests? *Journal of Statistics Education*, 17(2), 1-19.
- Harradine, A., Batanero, C. y Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 235–246). New York: Springer.
- Kadijevich, D., Kokol-Voljc, V. y Lavicza, Z. (2008). Towards a suitable designed instruction on statistical reasoning: Understanding sampling distribution with technology. En C. Batanero, G. Burrill, C. Reading y A. Rossman (Eds.), *Proceedings of the ICMI Study 18 Conference and IASE 2008 Round Table Conference*. Monterrey: International Statistical Institute. Disponible en https://www.stat.auckland.ac.nz/~iase/publications/rt08/T4P9_Kadijevich.pdf
- Moreno, A. (2018a). El razonamiento implicado en el estudio del muestreo. Aportes para la enseñanza de la estadística inferencial. *Acta Latinoamericana de Matemática Educativa*, 31(1), 757-763.
- Moreno, A. (2018b). The construction of the central ideas of the sampling through technology in university students. Trabajo presentado en el *10th. International Conference on Teaching Statistics "Looking back, looking forward"*, Kyoto, Japan: IASE.
- Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, 65(2), 123–155.
- Rubin, A., Hammerman, J. K. L., y Konold, C. (2006). Exploring informal inference with interactive visualization software. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town, South Africa: International Association for Statistics Education. Disponible en: [https://www.ime.usp.br/~abe/ICOTS7/Proceedings/PDFs/InvitedPapers/2D3 RUB I.pdf](https://www.ime.usp.br/~abe/ICOTS7/Proceedings/PDFs/InvitedPapers/2D3_RUB I.pdf)
- Saldanha, L. y Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257–270.
- Wild, C. J., Pfannkuch, M., Regan, M. y Horton, N. J. (2011). Towards more accessible conceptions of statistical inference. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(2), 247–295.
- Zieffler, A., Garfield, J. B., delMas, R., y Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7(2), 5-19. Disponible en: https://www.stat.auckland.ac.nz/~iase/serj/SERJ7%282%29_Zieffler.pdf