

## El sentido del muestreo

Carmen Batanero, M<sup>a</sup> Magdalena Gea y Nuria Begué  
(Universidad de Granada. España)

### 1. Introducción

En este trabajo analizamos las componentes que, a nuestro juicio, formarían el sentido del muestreo, tema que hemos elegido por ser el fundamento en que se apoya la inferencia estadística, que es hoy día un instrumento indispensable en la técnica, la ciencia y la gestión. Utilizamos, además, el muestreo en muchas actividades cotidianas; por ejemplo, cuando controlamos nuestra tensión arterial, a partir de algunas mediciones periódicas o estimamos el tiempo de realización de cierta tarea. En todas estas ocasiones partimos del análisis de muestras para construir nuestro conocimiento y realizar predicciones, ya que no podemos observar la totalidad del fenómeno que nos interesa.

Tanto Heitele (1975) como Burrill y Biehler (2011) describen el muestreo como idea estadística fundamental, al enlazar estadística y probabilidad y porque su estudio y la obtención de modelos de las distribuciones muestrales, contribuyeron al crecimiento de la estadística. Por otro lado, las operaciones combinatorias se definen mediante muestras. Sin embargo, Begué, Batanero, Ruiz y Gea (en prensa) describen numerosos sesgos en la comprensión de dicho concepto.

### 2. El sentido del muestreo

En este artículo retomamos dos de nuestros trabajos previos en la revista (Batanero, Díaz, Contreras y Roa, 2013; Gea, Batanero y Roa, 2014) en los que destacábamos la necesidad de reforzar el sentido estadístico de los estudiantes, que asumimos, se compone de unas actitudes y creencias adecuadas, la comprensión y capacidad de aplicación de las ideas estadísticas fundamentales y, el razonamiento estadístico, que ahora concretamos para el caso del muestreo.

#### 2.1. Actitudes y creencias

Di Martino y Zan (2015), entre otros, resaltan la importancia del componente emocional en la enseñanza de las matemáticas y la forma en que afecta al aprendizaje. Es importante, entonces, que el estudiante adquiera unas actitudes positivas hacia el muestreo y lo valore como instrumento de resolución de problemas, así como que adquiera una actitud crítica ante la información basada en muestreo, pues puede ocurrir que el procedimiento de muestreo utilizado no haya sido el adecuado. Además, se deben superar creencias erróneas sobre el tema, entre otras, las siguientes:

- *Insensibilidad al tamaño de la muestra* (Kahneman, Slovic y Tversky, 1982), es decir, asumir que una muestra, aunque sea pequeña, e independientemente de cómo se ha seleccionado, tendrá una distribución y unos estadísticos muy similares a los de la población, generalizando indebidamente la ley de los grandes números. Por ejemplo, al tomar una muestra pequeña de recién nacidos, esperar que la mitad exactamente sean niñas.



- *No se percibe la independencia de los elementos de la muestra*; por ello, al tomar una muestra, se esperan rachas cortas de elementos similares (Kahneman, Slovic y Tversky, 1982). En el ejemplo anterior, si se van tomando recién nacidos para formar la muestra, si los dos primeros son varones, se espera que el tercero sea una niña.
- *Ilusión de control* (Langer, 1982), también relacionado con la falta de percepción de la independencia. Consiste en creer que se puede controlar los resultados de un proceso aleatorio para obtener un resultado deseado. Aparece frecuentemente en los jugadores compulsivos y se refuerza cuando se incrementa la motivación (por ejemplo, si se espera un fuerte premio) y la persistencia (cuanto más se juega).
- *Falacia de la composición* (Chernoff y Russel, 2012), que consiste en transferir a un todo una propiedad que se cumple en una de sus partes. En nuestro ejemplo, ya que los elementos niño y niña son equiprobables, al tomar una muestra de cuatro recién nacidos se otorga la misma probabilidad a obtener 0, 1, 2, 3 o 4 niñas.

### 2.2. Ideas fundamentales en el muestreo

Un sentido adecuado del muestreo, supone la comprensión y capacidad de aplicación de los siguientes conceptos y propiedades:

- *Muestra y población*: Se requiere diferenciar entre población y muestra (subconjunto de la población) y comprender que diferentes muestras de la misma población varían, pero, a la vez, pueden tener elementos comunes. Es importante comprender las razones que nos llevan a tomar muestras, como falta de tiempo o recursos o, el hecho de que el estudio realizado con una muestra proporcione una información suficiente para nuestros fines.
- *Representatividad y variabilidad muestral*: Las muestras son semejantes entre sí, porque todas se han tomado de la misma población y, por tanto, se espera que sus resúmenes estadísticos (por ejemplo, media muestral) se asemejen a los de la población (media poblacional). Pero, al mismo tiempo las muestras varían, por lo que hay que esperar que diferentes muestras tengan diferentes resúmenes estadísticos (la media muestral es una variable). La variabilidad en las muestras depende de la variabilidad en la población, pero puede disminuirse si se aumenta el tamaño de la muestra (Saldahna y Thompson, 2002).
- *Tres tipos de distribución*: La distribución interviene en el muestreo con tres niveles de complejidad (Harradine, Batanero y Rossman, 2011): a) la distribución teórica de probabilidad, que modela una variable aleatoria tomada de una población; b) la distribución del conjunto de datos que constituye una muestra aleatoria simple; c) la distribución de un estadístico (por ejemplo, la media muestral) en todas las posibles muestras del mismo tamaño y condiciones (distribución muestral). La coordinación de estos tres niveles de distribución y su diferenciación, supone una gran dificultad conceptual para los estudiantes.
- *Estimación*: La estimación es el proceso de obtención de los valores de los parámetros a partir de una muestra de la población; por ejemplo, podemos estimar la media de altura en la población de chicos de una cierta edad ( $\mu$ ) a partir de la media de altura en una muestra de dicha población ( $\bar{x}$ ). Por supuesto, las conclusiones tendrán siempre un grado de incertidumbre, que la teoría del muestreo trata de estudiar.
- *Métodos de muestreo y sesgo*: Para poder estudiar el error en la estimación, el método de selección debe asegurar que podemos calcular la probabilidad de que cada elemento de la población forme parte de la muestra y que los elementos muestrales sean independientes entre sí. La primera forma de conseguirlo es utilizar un método de selección aleatoria, pero es posible también utilizar métodos equivalentes (como el muestreo sistemático) o métodos donde, aunque no todos los elementos de la población tienen la misma probabilidad de ser elegidos, podemos medir esta probabilidad, como el muestreo estratificado. Si no somos capaces de asegurar unas condiciones adecuadas, se produce un sesgo en el muestreo que se transfiere a las conclusiones, que no disminuye aunque aumente el tamaño de la muestra.

### 2.3. Razonamiento estadístico

El sentido del muestreo requiere también desarrollar en los estudiantes un adecuado razonamiento estadístico que incluye, según Wild y Pfannkuch (1999), los siguientes tipos de razonamiento, que aplicaremos a la noción de muestreo:

- *Reconocer la necesidad de los datos*: Es esencial reconocer las situaciones en que es útil el muestreo, el tipo de datos que deben formar parte de la muestra, de qué poblaciones y variables deben recogerse, así como el tamaño de la muestra y tipo de muestreo necesitado.
- *Percibir la variabilidad*: Es importante reconocer la variabilidad en la población, en las muestras de la misma, en la distribución del estadístico muestral y en los datos recogidos. Dos fines importantes de la enseñanza de la estadística es que los estudiantes perciban estas fuentes de variabilidad y manejen modelos que permitan controlarla y predecirla.
- *Razonar con modelos estadísticos*: Los estudiantes necesitan diferenciar los datos que forman parte de la muestra (realidad) y del modelo (población), posiblemente descrito por una distribución como la binomial o normal. Se debe fomentar tanto la competencia para generar muestras a partir de un modelo de población, como la de caracterizar el modelo de población y estimar sus parámetros a partir de una muestra.
- *Integrar la estadística y el contexto*: Se refiere a la capacidad de interpretar los resultados del muestreo desde el contexto en que se recogieron los datos, para saber qué dicen de la población. Esta competencia es crucial, pero muchos estudiantes que realizan los cálculos estadísticos sin problema, tienen dificultad en interpretar los resultados dentro del contexto.

### 3. ¿Cómo desarrollar el sentido del muestreo en los estudiantes?

La exposición realizada muestra que el desarrollo del sentido del muestreo requiere de un periodo largo de formación, por lo que debiera comenzarse su enseñanza desde el comienzo de la Educación Secundaria Obligatoria. Se debe comenzar por actividades de obtención de muestras, utilizando material manipulativo, como fichas o bolas en urnas, donde los estudiantes puedan observar la variabilidad muestral y el efecto del tamaño de la muestra sobre dicha variabilidad. Para corregir los sesgos y creencias erróneas de los estudiantes se pueden realizar experiencias de simulación, utilizando alguno de los recursos disponibles en Internet. La disponibilidad actual de tecnología hace posible dedicar el tiempo que previamente se invertía en tareas rutinarias y cálculo, para trabajar de forma más intuitiva y reforzar el sentido del muestreo. Por ejemplo, el applet de la colección de Allan Rosman y Beth Chance (<http://www.rossmanchance.com/applets/OneSample.html>) visualiza las tres distribuciones que intervienen en el muestreo: la de los datos de la muestra, la distribución de la población y la distribución muestral del estadístico. El applet permite cambiar el tipo de distribución, el estadístico que se estudia y el tamaño de la muestra y ver su efecto sobre la distribución muestral.

Será interesante trabajar con proyectos estadísticos, donde los estudiantes partan de una situación que les interese (por ejemplo, ¿tienen mayor número de pulsaciones por minuto los chicos o las chicas?), donde necesariamente hayan de recurrir al muestreo (hay que tomar una muestra de chicos y chicas; hay que decidir las condiciones en que se miden las pulsaciones como la hora del día o si se toman antes o después de hacer ejercicio; decidir el tamaño y forma de recoger la muestra, etc.). Se tendrá que llevar a cabo la recogida y análisis de datos y obtener una conclusión, que se debe interpretar en el contexto del problema. Todo ello reconociendo la incertidumbre de los resultados y, en lo posible, dando una medida del posible error en las conclusiones. Dependiendo del tipo de estudiante, esta investigación u otras similares se pueden llevar a cabo con mayor o menor grado de formalización, llegando incluso a la inferencia formal y la estimación mediante intervalos de confianza en los estudiantes de Bachillerato. Pero incluso realizándola de manera intuitiva, con un simple estudio exploratorio, permitirá ir incrementando el conocimiento y el sentido del muestreo en los estudiantes.



### Bibliografía

- Batanero, C., Díaz, C., Contreras, J. M. y Roa, R. (2013). El sentido estadístico y su desarrollo. *Números* [en línea], 83. Recuperado el 29 de noviembre de 2018, de <http://www.sineuton.org/numeros/>
- Begué, N., Batanero, C., Ruiz, K. y Gea, M. M. (en prensa). Understanding sampling: a summary of the research. *Boletín de Estadística e Investigación Operativa*.
- Burrill, G. y Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En Batanero, C., Burrill G. y Reading C. (eds.) *Teaching statistics in school mathematics. Challenges for teaching and teacher education - A joint ICMI/IASE study*, 57-69. Springer: Dordrecht.
- Chernoff, E. J. y Russell, G. L. (2012). The fallacy of composition: Prospective mathematics teachers' use of logical fallacies. *Canadian Journal for Science, Mathematics and Technology Education*, 12 (3), 259-271.
- Di Martino, P. y Zan, R. (2015). The construct of attitude in mathematics education. En Pepin, B. y Roesken-Winter, B. (eds.) *From beliefs to dynamic affect systems in mathematics education: Exploring a mosaic of relationships and interactions*, 51-72. Springer: New York.
- Gea, M., Batanero, C. y Roa, R. (2014). El sentido de la correlación y regresión. *Números* [en línea], 87. Recuperado el 29 de noviembre de 2018, de <http://www.sineuton.org/numeros/>
- Harradine, A., Batanero, C. y Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. En Batanero, C., Burrill G. y Reading C. (eds.) *Teaching statistics in school mathematics. Challenges for teaching and teacher education*, 235-246. Springer: The Netherlands.
- Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics*, 6, 187-205.
- Kahneman, D., Slovic, P. y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Langer, E.J. (1982). The illusion of control. En Kahneman, D., Slovic, P., y Tversky, A. (eds.) *Judgment under uncertainty: Heuristic and biases*, 231-238. Cambridge University Press: New York.
- Saldanha, L. y Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257-270.
- Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67 (3), 221-248.

**Carmen Batanero.** Facultad de Ciencias de la Educación, Universidad de Granada. Fue miembro del Comité Ejecutivo de ICMI (International Comisión on Mathematical Instruction) y Presidenta de IASE (International Association for Statistical Education). Ha coordinado varios congresos y proyectos de educación estadística.

Email: batanero@ugr.es

**María Magdalena Gea Serrano.** Facultad de Ciencias de la Educación, Universidad de Granada. Doctora en ciencias de la educación, licenciada en matemáticas y en ciencias estadísticas, investigación en didáctica de estadística, probabilidad y combinatoria. Actualmente coordina el grupo de investigación de la Sociedad Española de Investigación en Educación Matemática en dicha área de investigación.

Email: mmgea@ugr.es

**Nuria Begué Pedrosa.** Departamento de Didáctica de la Matemáticas, Universidad de Zaragoza. Graduada en Matemáticas por la Universidad de Zaragoza. Máster en Didáctica de la Matemática y estudiante de doctorado en Ciencias de la Educación en la Universidad de Granada.

Email: nbegue@correo.ugr.es