

*Investigación*

# Comprensión de la representatividad y variabilidad muestral por estudiantes de Educación Secundaria

Nuria Begué, Carmen Batanero, María Magdalena Gea y Pablo Beltrán-Pellicer

## Resumen

El objetivo del trabajo fue analizar la mejora con el curso escolar de la comprensión de la representatividad y variabilidad muestral en estudiantes de secundaria. Para ello se analizan las respuestas dadas por una muestra de 536 estudiantes de tres cursos (segundo y cuarto curso de Educación Secundaria Obligatoria y segundo curso de Bachillerato) a una tarea en que se piden cuatro valores de una distribución binomial. Se analizan la media y rango de dichos valores, que constituyen una muestra de cuatro resultados de la citada distribución, comparando con los valores teóricos de la distribución muestral de la media y el rango. Los resultados sugieren que los estudiantes tienen una percepción aceptable del valor esperado, pero que la percepción de la variabilidad es incorrecta y no mejora con la edad. Se observan también el sesgo de equiprobabilidad y la heurística de la representatividad en una parte de los estudiantes, mejorando los resultados según avanza el curso.

**Palabras clave:** muestreo, representatividad y variabilidad, comprensión, estudiantes de secundaria.

## **Understanding sampling representativeness and variability by secondary school students**

### **Abstract**

The objective of the work was to analyze the improvement with the school year of the understanding of the representativeness and sample variability in high school students. For this, we analyze the answers given by a sample of 536 students from three courses (second and fourth year of Secondary Education and second year of Baccalaureate) to a task in which four values of a binomial distribution are requested are analyzed. We analyze the average and the range of these values, which constitute a sample of four results of that distribution, comparing with the theoretical values of the sample distribution of the mean and range. The results suggest that students have a good perception of the expected value, but that the perception of variability is incorrect and does not improve with age. We also observe the equiprobability bias and the heuristic representation in a part of the students, improving the results in higher levels.

**Keywords:** sampling, representativeness and variability, understanding, high school students.

**Nombre y apellido del autor 1:** Nuria Begué

**Dirección completa:** Área de Didáctica de la Matemática, Universidad de Zaragoza, Facultad de Educación, C/ Pedro Cerbuna, 12, 50009 Zaragoza, España

**Lugar de trabajo:** Universidad de Zaragoza. España

**Teléfono de contacto:** 976 761 301

**Dirección electrónica:** nbegue@unizar.es

**Nombre y apellido del autor 2:** Carmen Batanero

**Dirección completa:** Departamento de Didáctica de la Matemática. Facultad de Ciencias de la Educación. Campus de Cartuja, s/n, 18071, Granada, España

**Lugar de trabajo:** Universidad de Granada. España

**Teléfono de contacto:** 958207169

**Dirección electrónica:** batanero@ugr.es

**Nombre y apellido del autor 3:** María Magdalena Gea

**Dirección completa:** Departamento de Didáctica de la Matemática. Facultad de Ciencias de la Educación. Campus de Cartuja, s/n, 18071, Granada, España.

**Lugar de trabajo:** Universidad de Granada. España

**Teléfono de contacto:** 958243952

**Dirección electrónica:** mmgea@ugr.es

**Nombre y apellido del autor 4:** Pablo Beltrán-Pellicer

**Dirección completa:** Área de Didáctica de la Matemática, Universidad de Zaragoza, Facultad de Educación, C/ Pedro Cerbuna, 12, 50009 Zaragoza, España.

**Lugar de trabajo:** Universidad de Zaragoza. España

**Teléfono de contacto:** 876554812

**Dirección electrónica:** pbeltran@unizar.es

## 1. Introducción

El muestreo es una idea básica en inferencia, según Heitele (1975) y juega un gran papel en el estudio de ciertos temas de probabilidad (leyes de los grandes números y enfoque frecuencial). Burrill y Biehler (2011) lo consideran como una de las ideas fundamentales en estadística por ser la base de la inferencia y del trabajo con la simulación, cuya utilización en el aula es recomendada para mejorar la comprensión de la probabilidad y la inferencia estadística (Eichler y Vogel, 2014).

Aunque la enseñanza del tema ha estado siempre presente en los currículos españoles de secundaria,

la nueva legislación (MECD, 2015) le proporciona mayor énfasis, tanto en sí mismo, como por su relación con el estudio del significado frecuencial, en el que la probabilidad se concibe como el límite teórico de la frecuencia relativa en una serie larga de experimentos. Dicho enfoque permite conectar la estadística y la probabilidad, pues la estimación de una probabilidad se debe realizar a partir de datos empíricos, lo que implica de hecho un proceso de muestreo (Batanero, 2005). Más concretamente, la Tabla 1 presenta los contenidos relacionados con el muestreo en el currículo básico (MECD, 2015).

Curso	Contenidos
1º y 2º Educación Secundaria Obligatoria (ESO).	Población e individuo. Muestra. Variables estadísticas. Fenómenos deterministas y aleatorios. Formulación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos y diseño de experiencias para su comprobación. Frecuencia relativa de un suceso y su aproximación a la probabilidad mediante la simulación o experimentación. (MECD, 2015, p. 413).
3º ESO (rama orientada a las enseñanzas académica y rama orientada a la enseñanza aplicada)	Fases y tareas de un estudio estadístico. Población, muestra. Variables estadísticas: cualitativas, discretas y continuas. Métodos de selección de una muestra estadística. Representatividad de una muestra. (MECD, 2015, p. 394).
Matemáticas Aplicadas a las Ciencias Sociales II (2º Bachillerato)	Población y muestra. Métodos de selección de una muestra. Tamaño y representatividad de una muestra. Estadística paramétrica. Parámetros de una población y estadísticos obtenidos a partir de una muestra. Estimación puntual. Media y desviación típica de la media muestral y de la proporción muestral. Distribución de la media muestral en una población normal. Distribución de la media muestral y de la proporción muestral en el caso de muestras grandes. Estimación por intervalos de confianza. Relación entre confianza, error y tamaño muestral. Intervalo de confianza para la media poblacional de una distribución normal con desviación típica conocida. Intervalo de confianza para la media poblacional de una distribución de modelo desconocido y para la proporción en el caso de muestras grandes. (MECD, 2015, p. 389)

**Tabla 1.**

*Contenidos relacionados con el muestreo en el currículo de secundaria.*

Las investigaciones previas sobre muestreo analizan la comprensión de diferentes propiedades e indican dificultades en la comprensión del tema (Begué, Batanero, Ruiz y Gea, 2019). No obstante, salvo en el trabajo de Gómez, Batanero y Contreras (2014), no se analiza la variabilidad de diferentes estimaciones del valor medio dadas por el mismo estudiante, ni la mejora de la comprensión de la variabilidad al avanzar el curso escolar. Dicho estudio lo abordamos en un trabajo anterior (Begué, Batanero y Gea, 2018), donde se observó un mejor desempeño en estudiantes de cuarto curso de Educación Secundaria Obligatoria respecto a los de segundo curso, pero todavía se identificaban muchas dificultades para generar muestras cuyo valor medio y rango fuera adecuado, siendo la identificación de la variabilidad más compleja para los estudiantes.

Para contribuir a completar estos estudios, el objetivo de la actual investigación es analizar si la mejora observada en el trabajo de Begué *et al.*, 2018 continúa en los estudiantes de Bachillerato y analizar la significatividad de las posibles diferencias. En lo que sigue se describe el marco teórico y los antecedentes, el método y los resultados del trabajo.

## 2. Marco teórico

La comprensión del muestreo requiere relacionar los conceptos de representatividad y variabilidad muestral (Saldahna y Thompson, 2002), el primero de los cuales indica que una muestra aleatoria de tamaño suficiente tendrá características parecidas a las correspondientes en la población, mientras que el segundo supone que las muestras pueden cambiar en su composición y características.

Además, en el muestreo se trabaja con tres tipos de distribuciones que los estudiantes deben comprender y relacionar (Harradine, Batanero y Rossman, 2011):

- La distribución teórica de probabilidad que modela los valores de una variable aleatoria en una población y depende de algún parámetro o característica estadística en dicha población.

- La distribución de datos en una muestra, donde podemos calcular un estadístico (que varía de muestra a muestra para estimar el valor del parámetro, que es constante, pero desconocido).

- La distribución muestral del estadístico o distribución de probabilidad de todos los valores que puede tomar el estadístico en el conjunto de las posibles muestras de la población de un tamaño dado.

## 3. Antecedentes

Las primeras investigaciones sobre el muestreo se relacionan con los trabajos sobre *heurísticas y sesgos*. Entre dichas heurísticas las siguientes son relevantes en nuestro trabajo:

- **La heurística de representatividad** (Tversky y Kahneman, 1982), que consiste en realizar una estimación, teniendo en cuenta únicamente la similitud que guarda la muestra con la población. Un sesgo asociado es la *insensibilidad al tamaño de la muestra*, que se manifiesta cuando no se tiene en cuenta el tamaño de la muestra para juzgar la variabilidad muestral. Otro sesgo (*falacia del jugador*) es considerar que el resultado de un experimento aleatorio afectará en la probabilidad de sucesos futuros. Si se supone que los siguientes resultados seguirán el patrón observado, se habla de *recencia positiva* y si se piensa que se compensarán los resultados futuros con los observados se denomina *recencia negativa*.

- **La heurística de la disponibilidad** (Tversky y Kahneman, 1974) consiste en estimar la probabilidad de un suceso basándose en la facilidad por encontrar ejemplos de situaciones similares. Un sesgo asociado es la *equiprobabilidad* (Lecoutre, 1992), mediante la cual se supone que los resultados de cualquier fenómeno aleatorio son igualmente probables. Este sesgo ha sido recientemente explicado por Chernoff y Russel (2012) mediante la *falacia de la composición*, que consiste en extender indebidamente la equiprobabilidad de los sucesos de un experimento aleatorio simple a cualquier muestra de resultados del experimento.

La tarea utilizada en este trabajo fue primeramente utilizada por Green (1983) en un estudio con niños ingleses (de 11 a 16 años), a los que proporciona los datos obtenidos en la realización de un experimento, donde 68 de 100 chinchetas caen con la punta hacia arriba y pregunta qué ocurriría al repetir el experimento. El autor esperaba que los niños diesen un resultado parecido, pero no idéntico. No obstante, los resultados revelaron que el 64% de los niños mostraron el sesgo de equiprobabilidad (Lecoutre, 1992), pues proporcionaron respuestas en las que la mitad de las chinchetas, aproximadamente, caerían hacia arriba, sin tener en cuenta la información frecuencial. Por otro lado, otros niños que mostraron una preferencia a obtener punta hacia arriba (correcta o parcialmente correcta), dieron una cantidad muy alejada de la esperada. Únicamente el 17% de la muestra dio una estimación correcta de la probabilidad en el experimento propuesto. Este mismo ítem fue propuesto por Cañizares (1997) a 253 niños españoles (de entre 10 a 14 años). La autora indica que el 64,1% de los alumnos muestran el sesgo de equiprobabilidad, mientras que solo el 15% responde correctamente, dando una estimación adecuada.

Orta y Sánchez (2013) realizan una investigación para explorar el razonamiento de una muestra de 65 estudiantes (14 años) sobre la noción de dispersión (variabilidad o variación). En concreto, los problemas que constituyen el cuestionario son de comparación de conjuntos de datos, los cuales presentan contextos de riesgo. Los resultados muestran la dificultad de los estudiantes para interpretar la dispersión, puesto que no toman una decisión racional a pesar de que identifican el conjunto de datos con mayor dispersión.

Sánchez, García y Medina (2014) consideran una muestra de estudiantes de secundaria (14-15 años) y Bachillerato (15-16 años) para identificar los rasgos característicos de los niveles de razonamiento frente a una situación-problema de probabilidad, en concreto, con la distribución binomial. Se establecen una serie de niveles estructurales para clasificar las respuestas siguiendo la taxonomía SOLO. Los

resultados revelan una mejora del razonamiento con la edad, aunque todavía persisten sesgos como el de equiprobabilidad (Lecoutre, 1992).

Más recientemente, Gómez, Batanero y Contreras (2014) proponen a una muestra de futuros profesores una adaptación del propuesto por Green (1983), que consiste en pensar cuatro resultados probables, si se repite el experimento. Los autores analizan la comprensión de la representatividad y variabilidad por parte de los participantes y encuentran que solo una tercera parte aproximadamente de la muestra tiene una intuición simultánea de las dos propiedades. Por otro lado, los participantes presentan diferentes sesgos como la equiprobabilidad, la heurística de la representatividad o piensan que no es posible realizar la predicción.

En nuestro trabajo anterior (Begué, Batanero y Gea, 2018) propusimos el ítem adaptado por Gómez *et al.* (2014) a estudiantes de Educación Secundaria Obligatoria, observando muchas dificultades, aunque una mejora según avanza el curso escolar. Retomamos ahora la misma tarea añadiendo a la muestra anterior un grupo de estudiantes de Bachillerato, y realizando un análisis estadístico complementario para observar las posibles diferencias al aumentar el curso escolar y si dichas diferencias son estadísticamente significativas.

#### 4. Metodología

Se trata de un estudio exploratorio de evaluación, básicamente cuantitativo puesto que nuestras conclusiones se deducen del análisis estadístico de los resultados que incluye tablas y gráficos, así como análisis de varianza.

La muestra está constituida por 536 alumnos, de los cuales 157 eran de 2º curso de ESO (13-14 años), 145 de 4º de ESO (15-16 años) y 234 de 2º de bachillerato (17-18 años). El número de grupos distintos de alumnos que participaron fue 9 grupos de 2º de ESO, 8 de 4º de ESO, de los cuales 6 grupos corresponden a la opción B, orienta a los alumnos que pretenden continuar el bachillerato.

En 2º bachillerato, participan cinco grupos, de los cuales 85 estudiantes cursan el bachillerato de Ciencias Sociales y 149 el bachillerato de Ciencias. Los participantes que configuran la muestra pertenecen a distintos centros educativos de la Comunidad Autónoma de Aragón. La forma en que se seleccionaron los estudiantes fue intencional, no obstante, al utilizar diferentes centros y contar con un alto número de estudiantes, se trató de aumentar la representatividad de la muestra. Los alumnos de 2º de ESO que forman parte de la muestra no habían recibido instrucción previa

sobre estadística, aparte de los conocimientos que pudieran haber obtenido en la educación primaria. Algunos de los alumnos de los grupos de 4º de ESO recibieron instrucción sobre probabilidad desde el significado clásico. En relación con los estudiantes de 2º de bachillerato, estos habían recibido instrucción previa sobre probabilidad y estadística. Los datos se recogieron por escrito, en la clase de matemáticas, como una actividad de la citada clase, explicándoles el fin de la evaluación y resolviendo las dudas sobre la forma de completarlo. En el cuestionario se propone a los estudiantes la siguiente tarea:

**Ítem 1.** Un profesor vacía sobre la mesa un paquete de 100 chinchetas obteniendo los siguientes resultados:

68 caen con la punta para arriba  y 32 caen hacia abajo 

Supongamos que el profesor pide a 4 niños repetir el experimento, lanzando las 100 chinchetas. Cada niño vacía una caja de 100 chinchetas y obtendrá algunas con la punta hacia arriba y otras con la punta hacia abajo. Escribe en la siguiente tabla un resultado que te parezca probable para cada niño:

<b>Daniel</b>	<b>Martín</b>	<b>Diana</b>	<b>María</b>
Punta arriba:	Punta arriba:	Punta arriba:	Punta arriba:
Punta abajo:	Punta abajo:	Punta abajo:	Punta abajo:

Este ítem, como se ha indicado, se ha tomado de Gómez *et al.* (2014), quien lo adaptó de otro anterior de Green (1983). El fenómeno aleatorio se corresponde con el lanzamiento de chinchetas y la distribución del número de chinchetas que caen con la punta hacia arriba se puede modelizar con un modelo binomial,  $B(n,p)$ . Estos datos permiten una estimación tanto del valor esperado en los 100 ensayos:  $np$ , como de su desviación típica:  $\sqrt{np(1-p)}$ . Con los datos de la tarea, la probabilidad estimada de  $p$  es 0,68 y el valor esperado de chinchetas hacia arriba de 68, con una desviación típica de 4,7. Por tanto, se puede considerar que el sujeto presenta una buena concepción o intuición del valor esperado si el valor medio de las cuatro respuestas que proporciona es cercano a 68. A este respecto, se considerarán cercanas o normativas aquellas res-

puestas cuyos valores, para el número de chinchetas que caen con la punta hacia arriba, se localicen dentro del intervalo  $[65,7-70,3]$ , que se calcula al determinar en la distribución muestral de la media el valor esperado (68) más/menos un error típico. Este es el intervalo de valores que contiene el 68% de las observaciones más probables en la distribución normal (la distribución muestral en este caso se aproxima bien por la normal). Si la media se sitúa en el intervalo  $[63,3-72,7]$ , que contiene el 95%, se considera aceptable. El alumno que proporcione estimaciones en este intervalo, debe también reconocer la ausencia de equiprobabilidad, por la asimetría física del dispositivo (puesto que la cabeza pesa más, es más probable que la chincheta caiga con la punta hacia arriba). Para analizar la comprensión de la variabilidad en diferentes muestras, se estudia la distribución mues-

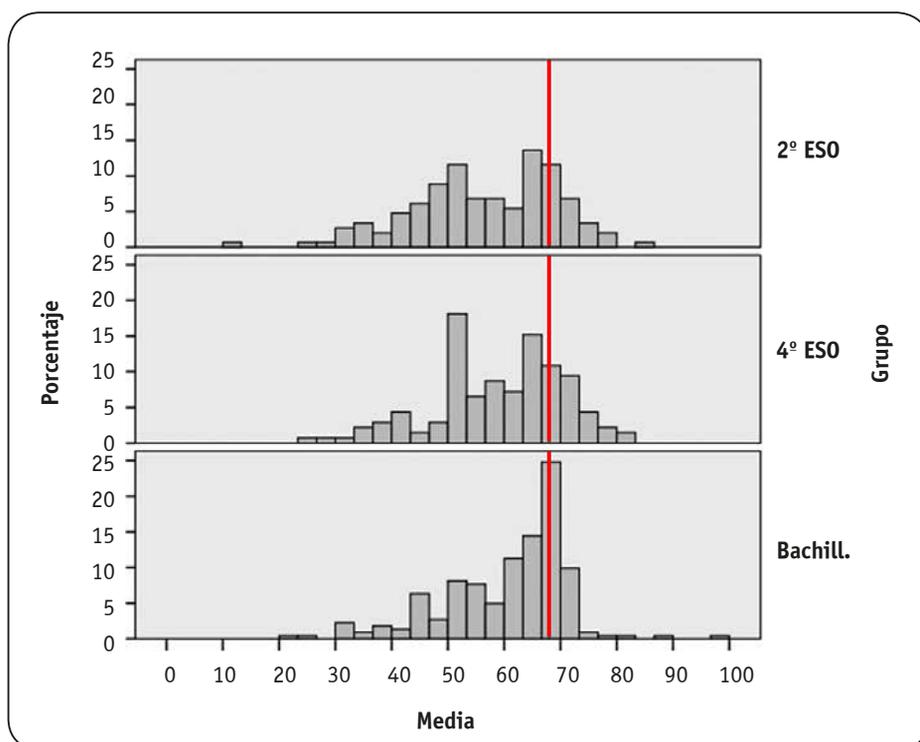
tral del rango de los cuatro valores proporcionados en la respuesta a la tarea, que no es normal, pero se ha simulado. En dicha distribución, el 95% de los valores centrales del rango varían entre 5 y 14, lo que consideramos como estimación normativa y, el 98% de los valores se sitúan entre 3 y 18, intervalo que incluye las respuestas de variabilidad aceptable.

## 5. Resultados y discusión

A continuación, presentamos en primer lugar los resultados que se refieren a la comprensión de la representatividad muestral, seguidos por los correspondientes al estudio de la variabilidad, finalizando un análisis de varianza para estudiar la significatividad de las diferencias encontradas.

### 5.1. Comprensión de la representatividad

Para deducir la comprensión de la representatividad muestral se calculó para cada estudiante la media de los cuatro valores encontrados y se analizó la distribución de las medias obtenidas de este modo para cada alumno, cuya distribución se representa, para cada uno de los cursos participantes en la Figura 1. Se obtuvo un valor medio empírico igual a 57,9, muy similar al obtenido en el trabajo de Gómez *et al.* (2014), que obtuvo un valor medio de 57,7 para dicha media, aún cuando el valor medio teórico es 68 (línea vertical señalada la Figura 1). Ello se explica porque una parte de los estudiantes proporcionan una estimación del número medio de chinchetas que caen hacia arriba sesgada hacia el valor 50, debido a que no tienen en cuenta la información frecuencial y consideran equiprobables los dos sucesos posibles del experimento.



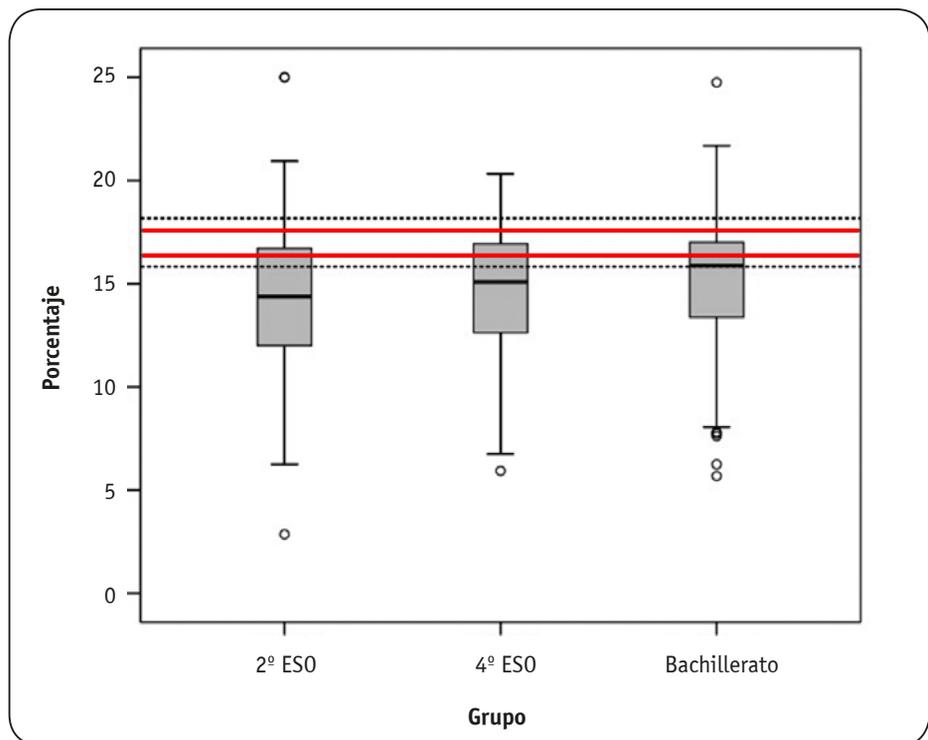
**Figura 1.**

*Distribución del número medio de 4 estimaciones de chinchetas con la punta hacia arriba.*

Sin embargo, al comparar los gráficos que corresponden a los diferentes grupos de estudiantes, se observan algunas diferencias. En los grupos de 2º y 4º curso de la ESO se observan claramente dos intervalos con alta frecuencia en la distribución, con valores medios de las cuatro estimaciones alrededor del 50 (estudiantes que caen en el sesgo de equiprobabilidad, descrito por Lecoutre, 1992) y alrededor del valor teórico (estudiantes que tienen una buena percepción de la representatividad muestral). Por el contrario, en los estudiantes de 2º curso de Bachillerato, la moda cae claramente alrededor del valor teórico, aunque todavía algunos proporcionan valores cuyo promedio está muy alejado de dicho valor.

Esta situación se visualiza mejor en los gráficos de caja presentados en la Figura 2, a los que hemos añadido en rojo, línea continua el intervalo de estimación normativa (68% de los valores teóricos

en la distribución muestral de la media de 4 valores de la distribución binomial) y en negro, línea punteada el intervalo correspondiente al 95% de dichos valores (estimación aceptable). Observamos que las medianas en las tres distribuciones de las medias de los alumnos quedan fuera del intervalo aceptable de estimación y justo en el borde inferior en el grupo de Bachillerato. Por tanto, la mitad de los alumnos de este grupo daría estimaciones cuyo valor medio se sitúa por debajo de lo aceptable y bastante más de la mitad en los otros dos grupos. Los valores obtenidos en esta distribución se detallan en la Tabla 2, donde se proporciona el porcentaje de estudiantes en cada grupo, según el valor medio de las cuatro estimaciones, donde observamos un crecimiento (hasta llegar al 20,8% en bachillerato) del porcentaje de estimaciones normativas e igualmente de aceptables (se añade desde un 17,8 a un 23,4% según grupo).



**Figura 2.**

*Distribución del número medio de 4 estimaciones de chinchetas con la punta hacia arriba.*

Valor medio de las cuatro estimaciones	2ºESO (n=157)	4ºESO (n=145)	Bachillerato (n=234)
Estimación normativa del valor esperado [65,7–70,3]	10,8	11,7	20,8
Estimación aceptable del valor esperado* [63,3–72,7]	17,8	21,4	23,4
Recencia negativa (menor que 45)	16,6	11,0	14,8
Recencia positiva (mayor que 72,7)	8,3	8,3	2,8
Equiprobabilidad [45–55]	26,8	23,4	17,7
Otros valores (55–63,3)	13,4	19,3	15,8
No completa	6,4	4,8	4,7

\*Fuera del intervalo normativo.

**Tabla 2.**

Porcentaje de estudiantes por grupo según valor medio en las cuatro estimaciones

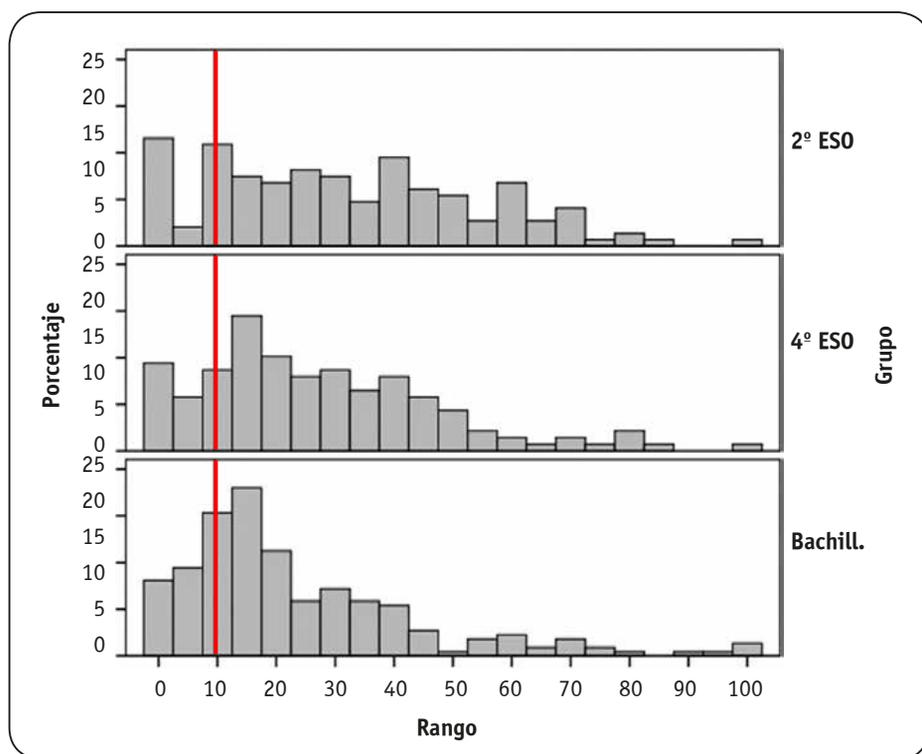
Por otro lado, observamos diferentes sesgos en los estudiantes. Los sesgos de equiprobabilidad y la heurística de la representatividad, también hallados en las muestras de sujetos que participan en la investigación de Gómez *et al.* (2014) y de Serrano (1996) se presentan con frecuencia apreciable, sobre todo el primero. Por un lado, algunos estudiantes presentan valores medios menores que 45 y hemos considerado que siguen una recencia negativa, pues como los datos indicaban un número de chinchetas con la punta hacia arriba igual a 68, muestran una tendencia a dar valores medios más bajos del 50%, es decir, consideran que el número de chinchetas que caen con la punta hacia arriba es menor, para compensar el resultado del profesor. Esta tendencia muestra la heurística de la representatividad (Tversky y Kahneman, 1982) y no desaparece en los alumnos mayores. Son pocos los que muestran la recencia positiva (tendencia a continuar o exagerar la línea del profesor).

Destacamos también los que dan valores medios muy próximos al 50% que indica el sesgo de equiprobabilidad, aunque el porcentaje disminuye con el curso escolar. Por último, en otros valores se engloban las respuestas de los alumnos que son incorrectas por diferentes razones y hay que añadir los que no responden al ítem.

## 5.2. Comprensión de la variabilidad muestral

Para analizar la comprensión de la variabilidad muestral se calculó el rango de las cuatro estimaciones dadas por cada estudiante, cuya distribución, según grupo, se muestra en la Figura 3, donde de nuevo hemos marcado con una línea vertical la media teórica de la distribución muestral de rangos (9,6) que es mucho menor que la media observada en la muestra (23,7) lo que indica que los estudiantes supone que la variabilidad muestral es mucho mayor que la que corresponde al experimento estudiado (distribución binomial, con tamaño de muestra igual a 100).

Observamos que la mayoría de los estudiantes produce una variabilidad excesiva (superior al valor teórico, marcado con una línea en el gráfico). Por tanto, son muchos los estudiantes que no comprenden el efecto del tamaño de muestra sobre la variabilidad del muestreo, en línea con la investigación de Serrano (1996). También en este caso se muestra una mejora al avanzar el curso, ya que la distribución se empieza a concentrar en los cursos superiores alrededor del valor teórico, aunque todavía con demasiada dispersión y con un sesgo a la derecha de dicho valor teórico.

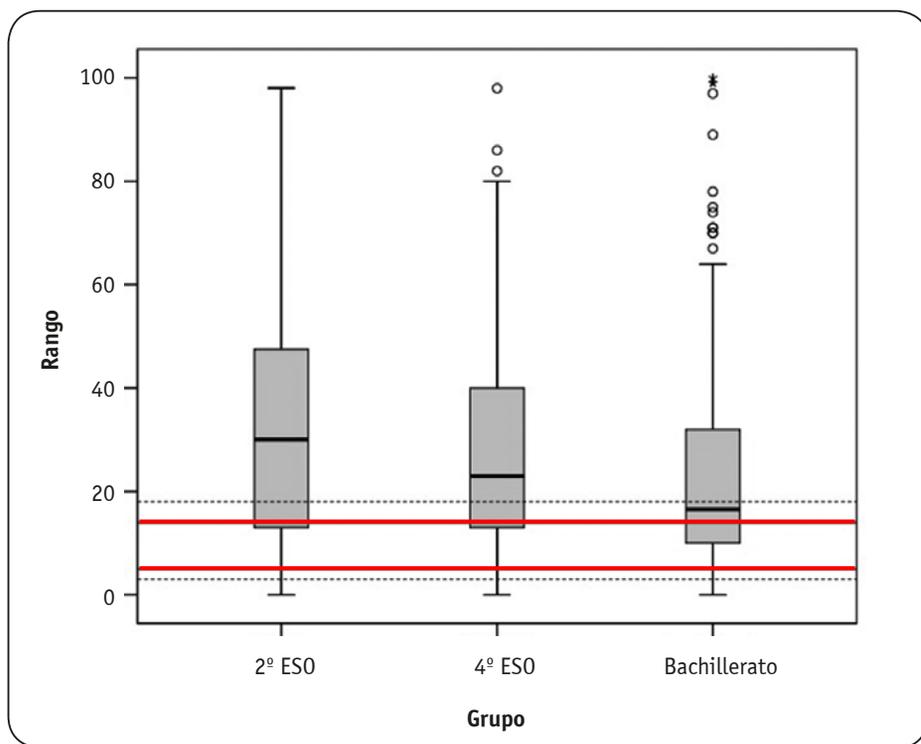


**Figura 3.**  
Distribución del rango de 4 estimaciones de chinchetas con la punta hacia arriba.

Esta situación se visualiza mejor en los gráficos de caja presentados en la Figura 4, a los que hemos añadido en rojo, línea continua el intervalo de estimación normativa (68% de los valores teóricos en la distribución muestral del rango de 4 valores de la distribución binomial) y en negro, línea punteada el intervalo correspondiente al 95% de dichos valores (estimación aceptable).

Observamos que la mayor parte de la distribución de rangos en los valores dados por los alumnos queda fuera del intervalo aceptable de estimación, debido a la variabilidad excesiva, aunque la proporción de los que quedan fuera disminuye con el curso, quedando algo menor que la mitad de la muestra en los estudiantes de bachillerato, que, por tanto, parecen tener mejor percepción de la variabilidad muestral. Aparecen algunos valores atípicos, que se corresponden con respuestas cuyos valores para el número de chinchetas se sitúa en los extremos del

rango de respuestas posibles, como se observa en las siguientes respuestas de los estudiantes (donde la letra A significa alumno): A61 y A135: (100, 100, 100, 100), A24: (22, 0, 8, 16) y A216: (20, 40, 20, 15), siendo A61 el alumno 61 que responde al cuestionario. Como señala Gómez *et al.* (2014), estos dos últimos resultados pueden indicar una creencia en la compensación entre los resultados, mostrando la heurística de la representatividad ya comentada. Estos gráficos se completan con la Tabla 3, en que clasificamos a los estudiantes según el rango de los cuatro valores aportados. En concreto, son pocos los que se sitúan en el intervalo de estimación normativa, creciendo el porcentaje por curso, y también creciendo el porcentaje de los que, fuera del intervalo anterior, proporcionan una estimación aceptable. En conjunto entre los dos intervalos se pasa del 20,3% en 2º de ESO al 40,7% en bachillerato, lo que es una mejora notable.



**Figura 4.**

*Distribución del rango de 4 estimaciones de chinchetas con la punta hacia arriba.*

Rango de las cuatro estimaciones	2ºESO (n=157)	4ºESO (n=145)	Bachillerato (n=234)
Estimación normativa [5-14]	12,7	17,9	19,5
Estimación aceptable [3-18]*	7,6	11,7	21,2
Estimación de una variabilidad muestral excesiva >18	62,4	56,6	42,8
Alta concentración <3	10,8	9	11,8
No completa	6,4	4,8	4,7

\*Fuera del intervalo normativo.

**Tabla 3.**

*Porcentaje de estudiantes por grupo según intervalo en que se sitúa el rango de las cuatro estimaciones.*

No obstante, el mayor porcentaje en todos los grupos es el de los estudiantes que proporcionan respuestas con una dispersión muy alta (mayor que 18), disminuyendo el porcentaje con el curso. Por tanto, estos estudiantes no comprenden que el tamaño de la muestra de la distribución binomial de partida reduce la variabilidad muestral. Por ello, dichos alumnos solo alcanzan uno de los niveles inferiores de razonamiento sobre muestreo (idiosincrático o de transición) según Moreno y Vallecillos (2001), puesto que no reconocen el efecto del tamaño de la muestra sobre la variabilidad. En Méndez (1991) y Serrano (1996) se obtiene la presencia de este error en sus investigaciones.

Además, una pequeña parte de los alumnos proporcionan datos con apenas variabilidad (pues el rango de las cuatro estimaciones dadas es menor que 3). Entre ellos, se incluye un 3,9% de respuestas en las que existe ausencia de variabilidad total; es decir, escriben la cuaterna (50, 50, 50, 50), respuesta que también aparece en el trabajo de Gómez *et al.* (2014). Por tanto, se repiten las dificultades con la idea de variabilidad observada en las investigaciones de Orta y Sánchez (2013) y Sánchez *et al.* (2014).

### 5.3. Resultados del análisis de varianza

Para comprobar si las diferencias observadas en los valores medios y rango en los tres grupos son estadísticamente significativas se ha realizado la prueba del análisis de varianza tomando el grupo de estudiantes como variable independiente. Los resultados se presentan en la Tabla 4, donde observamos que la diferencia de medias no es estadísticamente significativa para el caso de los valores medios y sí es muy significativa la diferencia de rangos, cuyo valor *p* es menor que 0,05. En conclusión, no tenemos motivos para suponer que los diferentes grupos de estudiantes muestran una comprensión diferente de la relación entre el valor esperado y el teórico en la tarea propuesta. Por el contrario, sí que se identifica que algunos grupos entienden mejor la variabilidad del muestreo en muestras grandes.

Desde el interés por comprobar qué grupos presentan una diferencia en la estimación del rango, se realizan las pruebas de Tukey de diferencia de medias, recomendadas una vez que el análisis de varianza da un valor significativo. Los resultados se resumen en la Tabla 5, donde solamente se obtiene diferencia estadísticamente significativa entre bachillerato y el segundo curso de ESO en el rango, pero no con cuarto curso.

		Suma de cuadrados	g.l.	Media cuadrática	F	Sig.
<b>I1Media</b>	<b>Inter-grupos</b>	290,111	2	145,055	,672	,511
	<b>Intra-grupos</b>	111378,865	516	215,851		
	<b>Total</b>	111668,976	518			
<b>I1Rango</b>	<b>Inter-grupos</b>	8198,057	2	4099,028	9,095	,000
	<b>Intra-grupos</b>	234348,303	520	450,670		
	<b>Total</b>	242546,359	522			

**Tabla 4.**

*Resultados del análisis de varianza.*

	Prueba t para la igualdad de medias				Intervalo de confianza de la diferencia (95%)	
	g.l.	Diferencia de medias	Error típ. diferencia	Sig. bilateral	E. Inferior	E. Superior
Diferencia con 2º	283	-9,517	2,234	,000	-14,77	-4,27
Diferencia con 4º	287	-4,125	2,258	,162	-9,43	1,18

**Tabla 5.**

*Prueba post-hoc de diferencias de medias para el rango de las cuatro estimaciones.*

## 6. Conclusiones e implicaciones para la enseñanza

Los resultados obtenidos nos llevan a reflexionar sobre el grado de consecución de los objetivos fijados en los documentos curriculares (MECD, 2015), donde se debiera poner más énfasis la introducción de las ideas necesarias para la comprensión adecuada y profunda del muestreo. Aunque los datos se recogieron el primer año de la aplicación de la Ley Orgánica de Mejora de la Calidad Educativa (LOMCE), los contenidos de probabilidad del Decreto de Enseñanzas Mínimas (MEC, 2007) eran casi idénticos, por lo que en aquellos grupos de alumnos que no recibieron enseñanza del tema, la única justificación posible es la falta de tiempo.

El análisis de los datos muestra una gran proporción de respuestas incorrectas en la estimación de los resultados a la tarea propuesta y también que los alumnos de la muestra presentan tanto el sesgo de equiprobabilidad como los derivados de la heurística de la representatividad. Esta última se identifica en aquellos casos en los que el alumno trata de compensar los resultados, proporcionando muestras en las que el número de chinchetas que caen con la punta hacia abajo sea mayor, al igual que ocurrió en el estudio de Gómez *et al.* (2014). Por otro lado, la mayoría de los alumnos tiene una comprensión insuficiente de la variabilidad intrínseca al proceso de muestreo, que se ha puesto de

manifiesto en el ítem propuesto para este estudio con muestras cuyo tamaño es grande ( $n=100$ ). En este ítem, la mayoría de las respuestas de los estudiantes presentan muestras cuyos valores quedan caracterizados por una variabilidad excesiva. Por tanto, se confirma una mayor dificultad en la idea de variabilidad, como ocurre también en la investigación de Orta y Sánchez (2013) en problemas relacionados con el contexto de riesgo y en la de Sánchez, García y Medina (2014) sobre la distribución binomial. También encontramos la presencia de respuestas en la que se identifica una ausencia total de la variabilidad.

A pesar de que se esperaba que los alumnos de 4º ESO obtuvieran mejores resultados, debido a factores tales como la mayor edad y su conocimiento, el análisis de las respuestas apenas revela diferencias con los de 2º curso. En síntesis, los resultados ponen en relieve la desconexión existente entre los contenidos descritos en los documentos curriculares y el conocimiento que los estudiantes muestran sobre dichos contenidos matemáticos. Además, los resultados obtenidos coinciden con las investigaciones previas, por tanto, se subraya la necesidad de iniciar una instrucción sobre los conceptos asociados al muestreo con la finalidad de favorecer su comprensión.

En este sentido, el proceso de enseñanza y aprendizaje debería comenzar desde los primeros cursos de la secundaria, de modo que se fortalezcan y desarrollen de manera gradual los contenidos asociados con el bloque de estadística y probabilidad. En particular, el estudio sugiere la necesidad de conectar el significado frecuencial de la probabilidad con el significado clásico, ofreciendo a los estudiantes oportunidades de experimentar situaciones donde

la regla de Laplace no puede aplicarse. Es el caso del experimento de chinchetas que hemos utilizado en este trabajo. Por otro lado, se requieren nuevas investigaciones que no solamente se preocupen en la identificación de sesgos o dificultades de los alumnos, sino también el diseño de material que favorezca una mejora del proceso educativo tanto al alumno como al docente.

**Agradecimientos:** Proyecto PID2019-105601GB-I00 (MICIN), Grupo FQM126 (Junta de Andalucía) y Grupo S36\_17D Investigación en Educación Matemática (Gobierno de Aragón y Fondo Social Europeo).

### Referencias bibliográficas

- Batanero, C. (2005).** Significados de la probabilidad en la educación secundaria. *Relime*, 8(3), 247-264.
- 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. DOI: <https://doi.org/10.5565/rev/ensciencias.2256>.
- Begué, N., Batanero, C., Ruiz, K. y Gea, M.M. (2019).** Understanding sampling: a summary of the research. *BEIO*, 35(1), 49-78.
- Burrill, G. y Biehler, R. (2011).** Fundamental statistical ideas in the school curriculum and in training teachers. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education* (pp. 57-69). Dordrecht, Holanda: Springer.
- 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. DOI: <http://doi.org/10.1080/14926156.2012.704128>
- Cañizares, M.J. (1997).** *Influencia del razonamiento proporcional y combinatorio y de creencias subjetivas en las intuiciones probabilísticas primarias*. Tesis doctoral. Universidad de Granada.
- Eichler, A., y Vogel, M. (2014).** Three approaches for modelling situations with randomness. En E. J. Chernoff y B. Sriraman (Eds.), *Probabilistic thinking, presenting plural perspectives* (pp. 75-99). Dordrecht, Holanda: Springer
- Gómez, E., Batanero, C. y Contreras, C. (2014).** Conocimiento matemático de futuros profesores para la enseñanza de la probabilidad desde el enfoque frecuencial. *Bolema*, 28(48), 209-229. DOI: <https://doi.org/10.1590/1980-4415v28n48a11>.
- Green, D.R. (1983).** A survey of probabilistic concepts in 3000 pupils aged 11-16 years. En D. R. Grey et al. (Eds.), *Proceedings of the First International Conference on Teaching Statistics* (Vol.2, pp. 766-783). Universidad de Sheffield: Teaching Statistics Trust.
- 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). Dordrecht, Holanda: Springer. [https://doi.org/10.1007/978-94-007-1131-0\\_24](https://doi.org/10.1007/978-94-007-1131-0_24)

- Heitele, D. (1975).** An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics*, 6(2), 187-205.
- Lecoutre, M.P. (1992).** Cognitive models and problem spaces in «purely random» situations. *Educational Studies in Mathematics*, 23, 557-568.
- MEC (2007).** *Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria*. Madrid: Ministerio de Educación y Cultura.
- MECD (2015).** *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato*. Madrid: Autor.
- Méndez, H. (1991).** *Understanding the central limit theorem*. (Tesis Doctoral). Universidad de California. UMI 6369
- Moreno, A. y Vallecillos, A. (2001).** Exploratory study on inferential' concepts learning in secondary level in Spain. En M. van der Heuvel (Ed.), *Proceedings of the 25th Conference of the International Group of the Psychology of Mathematics Education (PME)* (p. 343). Utrech: Freudenthal Institute and Utrecht University
- Orta, J. A., y Sánchez, E. (2013).** Interpretación de la dispersión de datos en contexto de riesgo por estudiantes de secundaria. *Investigación en Educación Matemática XVII*, 422-430.
- Saldanha, L., & Thompson, P. (2002).** Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257-270.
- Sánchez, E., García, J. y Medina, M. (2014).** Niveles de razonamiento y abstracción de estudiantes de secundaria y bachillerato en una situación-problema de probabilidad. *Avances de Investigación en Educación Matemática*, 6, 5-23. DOI: <https://doi.org/10.35763/aiem.v1i6.90>.
- Serrano, L. (1996).** *Significados institucionales y personales de objetos matemáticos ligados a la aproximación frecuencial de la enseñanza de la probabilidad*. Tesis doctoral. Universidad de Granada.
- Tversky, A. y Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131. DOI: <http://doi.org/10.1126/science.185.4157.1124>.
- Tversky, A. y Kahneman, D. (1982).** Judgments of and by representativeness. En D. Kahneman, P. Slovic y A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 117-128). New York: Cambridge University Press.