

Comprensión de elementos básicos de muestreo en alumnos de
educación secundaria obligatoria

Nuria Begué Pedrosa



TRABAJO DE FIN DE MÁSTER

Tutora: Carmen Batanero Bernabeu

Universidad de Granada

Departamento de Didáctica de la Matemática

2016

AGRADECIMIENTOS

Por un lado, me gustaría dedicar estas líneas para agradecer el trabajo de mi tutora, Carmen Batanero, la cual no ha sido solamente una guía durante la elaboración del mismo, sino que ha supuesto un apoyo gracias a sus consejos y comentarios. Por tanto, destacar su esfuerzo dedicado a la elaboración del presente trabajo.

Por otro lado, quiero agradecer tanto al cuerpo docente como a los alumnos que han participado en el estudio por haber compartido su tiempo y disposición, sin los cuales no hubiera sido posible la realización del mismo.

Además, me gustaría dedicar unas líneas a todas aquellas personas que desde un segundo plano han permitido que el camino haya sido más amable, ofreciendo su apoyo y confianza en mi trabajo. En especial, a Javier, por su apoyo incondicional durante todo este tiempo.

Finalmente, este trabajo forma parte del proyecto EDU 2013-41141-P (MEC) y Grupo FQM-126 (Junta de Andalucía).

COMPRENSIÓN DE ELEMENTOS BÁSICOS DE MUESTREO EN ALUMNOS DE EDUCACIÓN SECUNDARIA OBLIGATORIA

INTRODUCCIÓN	1
CAPÍTULO 1. PLANTEAMIENTO DE LA INVESTIGACIÓN	5
1.1. Introducción	5
1.2. Justificación del tema	6
1.3. Algunos elementos teóricos	7
1.4. Objetivos e hipótesis de la investigación	8
1.5. Marco curricular	10
1.5.1. El muestreo en los Decretos de Enseñanza Mínima	11
1.5.2. El muestreo en el Currículo Básico	14
1.5.3. Conclusiones	16
1.6. Contenidos sobre muestreo que se consideran en el trabajo	17
 CAPÍTULO 2. ANTECEDENTES	 19
2.1. Introducción	19
2.2. Niveles de uso de los conceptos de inferencia	20
2.3. Comprensión de las propiedades de muestreo	23
2.4. Comprensión del muestreo	24
2.5. Comprensión de la distribución muestral	25
2.6. Comprensión de la probabilidad, desde el punto de vista frecuencial	28
2.7. Conclusiones	31
 CAPÍTULO 3. ESTUDIO DE EVALUACIÓN	 33
3.1. Introducción	33
3.2. Contexto educativo	33
3.3. Metodología	34
3.3.1. Descripción de la muestra	35
3.3.2. Método de recogida de datos	35
3.4. Análisis a priori del cuestionario y conocimientos evaluados	36
3.4.1. Análisis del ítem 1	37
3.4.2. Análisis del ítem 2	38
3.4.3. Análisis del ítem 3	39

3.4.4. Análisis del ítem 4	39
3.5. Resultados	40
3.5.1. Resultados en el ítem 1	41
3.5.2. Resultados en el ítem 2	46
3.5.3. Resultados en el ítem 3	50
3.5.4. Resultados en el ítem 4	55
3.6. Conclusiones del análisis de los datos	58
3.6.1. Síntesis de resultados	59
3.6.2. Comparación de resultados según el tamaño de muestra	61
3.6.3. Comparación de resultados en experimentos con sucesos equiprobables o no	62
CAPÍTULO 4. CONCLUSIONES	63
4.1. Introducción	63
4.2. Conclusiones respecto a los objetivos	63
4.3. Conclusiones respecto a las hipótesis iniciales	64
4.4. Limitaciones del trabajo	66
4.5. Futuras líneas de investigación	67
4.6. Implicaciones para la enseñanza	68
REFERENCIAS	69
ANEXO I. ESTUDIO DE CASOS EXTREMOS	74

INTRODUCCIÓN

La inferencia estadística es hoy día una herramienta básica en muchas de las actividades humanas como, la ciencia, la gestión y la política. Aunque su estudio formal se retrasa al Bachillerato (y sólo para los alumnos de Ciencias Sociales), encontramos en la Educación Secundaria Obligatoria algunas ideas elementales sobre muestreo, cuya comprensión es básica para el estudio posterior de la inferencia estadística.

Se puede pensar en la inferencia como una corrección de herramientas que nos ayudan a generalizar las conclusiones obtenidas de una muestra a la población de donde se ha recogido. Según Batanero y Díaz (2015), las técnicas de inferencia se desarrollaron para fundamentar las formas de obtener un conocimiento general a partir del análisis de casos particulares, las cuales adquieren una gran importancia en las ciencias empíricas. Por tanto, la elección del muestreo como tema de investigación se fundamenta en que constituye un contenido fundamental para el aprendizaje de la Estadística (Burril y Biehler, 2011). De hecho, gran parte de nuestro conocimiento está basado en información obtenida a partir de muestras.

Aunque aparentemente este tema sea sencillo, la investigación previa, tanto con niños como con sujetos adultos, indica la existencia de creencias erróneas sobre el muestreo, que llevan a la toma de decisiones equivocadas en diversas situaciones de la vida diaria o del trabajo. De ello se deduce la necesidad de analizar mediante la investigación dichos razonamientos y ver la forma de mejorarlo con la enseñanza.

En este trabajo nos centramos específicamente en analizar la comprensión de algunas ideas elementales de muestreo por parte de una muestra de estudiantes de Educación Secundaria Obligatoria. Más concretamente, evaluamos su comprensión acerca de la variabilidad en el muestreo y del efecto del tamaño de la muestra sobre dicha variabilidad. También analizamos la comprensión de la relación entre el valor de una proporción en la población y la frecuencia relativa esperada en muestras tomadas de dicha población. La consideración de estos puntos conduce a la elaboración de un trabajo de investigación. Por tanto, este documento supone la Memoria del trabajo y, su propósito es presentar de manera sintetizada el estudio realizado. A continuación, se presenta una síntesis de cada uno de los capítulos que configuran dicha Memoria.

En primer lugar, el interés recae en justificar la importancia del tópico elegido que conduce a la identificación tanto de los objetivos como de las hipótesis que enmarcan el

problema de investigación planteado. Además, en este primer Capítulo 1 se considera el marco curricular con la finalidad de identificar la presencia de los conceptos, que son objeto de estudio, en los currículos oficiales.

Posteriormente, el trabajo se centra en revisar la literatura existente relacionada con el tópico que nos ocupa. Por tanto, el Capítulo 2 presenta una síntesis de aquellas investigaciones que se han dirigido hacia el análisis de la comprensión de las propiedades del muestreo, la distribución muestral y la comprensión de la probabilidad desde su significado frecuencial.

Este análisis previo guía en el diseño del instrumento de evaluación descrito de manera detallada en el tercer capítulo. Así mismo, este capítulo contempla la descripción detallada del contexto escolar, la metodología seguida y la muestra. Previamente al análisis de los datos constituidos por las respuestas dadas por los estudiantes, se analiza el cuestionario desde la reflexión sobre la respuesta esperada para cada uno de los ítems que lo constituye. Luego, se lleva a cabo un análisis de las respuestas para cada ítem. Además, este estudio se complementa con un análisis comparativo de las respuestas según el grupo participante. Por último, las características de cada ítem conducen a un análisis de los resultados en relación a las mismas.

Finalmente, el Capítulo 4 implica la discusión de los resultados, lo que supone su reflexión en relación con los objetivos e hipótesis descritas. Por otro lado, resulta imperante la cavilación acerca de las limitaciones que presenta el trabajo, junto con las posibles vías de investigación. Por último, se adjuntan las referencias bibliográficas que han sido consideradas para la elaboración del presente trabajo.

Este trabajo me ha permitido meditar sobre aquellos conceptos o ideas que han sido transmitidas a lo largo de las asignaturas que constituyen el Máster. En particular, su elaboración me ha permitido comprender y valorar la tarea investigadora que se desarrolla desde esta área, la Didáctica de la Matemática. Además, considero que la reflexión de los diferentes aspectos que configuran una investigación permite al estudiante conectar de una manera racional y reflexiva el trabajo que ha estado desarrollando a lo largo del curso, siendo una oportunidad de aprendizaje culmen para estos estudios.

Por otro lado, desde la perspectiva de docente, el tipo de estudio realizado pone en relieve la existencia de las intuiciones o concepciones previas que los estudiantes presentan sobre un contenido matemático concreto. En particular, la identificación de errores en las mismas supone un argumento suficiente para reflexionar sobre un proceso de enseñanza y aprendizaje que permita un progreso adecuado en la comprensión de los conceptos, con la finalidad de que no supongan un obstáculo en el aprendizaje del alumno.

CAPÍTULO 1.

PLANTEAMIENTO DE LA INVESTIGACIÓN

1.1. INTRODUCCIÓN

El interés por la enseñanza de la Estadística ha aumentado en los últimos años debido a la necesidad reclamada por la UNESCO y otras instituciones de ofrecer una formación estadística general al ciudadano, llevando a la idea de cultura estadística (Arteaga, Batanero, Cañadas y Contreras, 2011). Este reconocimiento ha trascendido a los planes de educación de diversos países, donde se observa el esfuerzo realizado en el diseño de currículos y recursos didácticos específicos. Batanero, Godino, Green, Holmes y Vallecillos (1994) citaron, como ejemplos de diseño curricular, el trabajo realizado por Holmes y cols. (1980) para el Schools Council Project on Statistical Education en Inglaterra y el libro sobre la probabilidad (Godino, Batanero y Cañizares, 1987). Más recientemente encontramos, por ejemplo, el libro de Watson, Beswick y Brown (2012) o las recomendaciones sobre la enseñanza incluidas en los diferentes capítulos del libro de Jones (2006).

Batanero y Godino (2005) muestran el progreso que ha experimentado la educación estadística, para lo cual presentan una síntesis de los avances realizados desde la propia Estadística, la Psicología y la Didáctica de la Matemática. En relación con el primero de los campos considerados, los autores resumen la tendencia creciente a incluir la enseñanza de la Estadística en todos los niveles educativos, gracias a las acciones impulsadas por el Instituto Internacional de Estadística (ISI). En este sentido, El ISI estableció en 1885 un Comité de Educación respaldado por la UNESCO, y uno de sus objetivos fue la creación de estudios reglados que formaran estadísticos, creando los Centros Internacionales de Educación Estadística, que también se han ocupado del diseño de materiales didácticos. Además, desde el ISI se impulsó la puesta en marcha de conferencias internacionales sobre la enseñanza de la Estadística, la novena de las cuales se ha celebrado en Estados Unidos en 2014. Por otro lado, en 1992 aparece el International Association for Statistical Education (IASE) que actualmente se encarga de organizar estos congresos.

Otro indicador del creciente interés hacia la enseñanza de la Estadística es la existencia de revistas específicas, entre las que destacan: *Journal of Statistics*

Education, Statistics Education Research Journal, Teaching Statistics, Statistique et enseignement, Induzioni y Stochastik in der Schule.

1.2. JUSTIFICACIÓN DEL TEMA

La inferencia estadística es una herramienta esencial, que forma parte de muchas ramas de la actividad humana. Este hecho explica que la enseñanza de este contenido esté presente en la Universidad y la formación profesional, así como en alguna modalidad de Bachillerato (Batanero y Díaz, 2015). A pesar de ser uno de los temas más enseñados, es también uno de los que peor se comprenden y aplican. Según Batanero (2013), ello se debe a la gran cantidad de conceptos y procedimientos que intervienen en la inferencia y al corto tiempo que se dedica a su enseñanza.

Como el muestreo es la base de la inferencia, es importante que los alumnos logren una comprensión suficiente del mismo, antes de continuar con el estudio de otros temas como el contraste de hipótesis y el intervalo de confianza, pues de lo contrario, los errores de comprensión del muestreo van a proyectarse en los contenidos posteriores. Heitele (1975) incluyó el concepto de muestreo en su lista de diez ideas estocásticas fundamentales, debido a su presencia en multitud de situaciones aleatorias, en las que aparece implícitamente esta idea. Además, la formalización de este concepto supuso un gran avance en el desarrollo de la estadística como ciencia. Por otro lado, según este autor, la enseñanza de este contenido matemático se puede iniciar desde edades tempranas con diversos grados de formalización.

Por otro lado, la comprensión del muestreo no sólo es necesaria en inferencia, sino en el día a día. Además de establecer un puente entre la estadística y la probabilidad. Su importancia es debida a que todo nuestro conocimiento y juicios sobre el mundo o las personas está basado en el muestreo, ya que, usualmente, sólo podemos estudiar u observar una parcela de la realidad en la que estamos interesados. Actualmente, los resultados obtenidos mediante muestreo en encuestas o investigaciones en campos que nos afectan (medicina, política, educación, consumo, etc.) se publican cada vez con mayor frecuencia en los medios de comunicación. Estas noticias, o los mismos estudios, a veces contienen sesgos no asumidos o intencionados, con fines comerciales o de otro tipo. Entonces, es relevante que los estudiantes aprendan a razonar sobre el muestreo como parte de su cultura estadística.

1.3. ALGUNOS ELEMENTOS TEÓRICOS

El marco teórico que se ha seguido para enmarcar las ideas teóricas que aparecen a lo largo del estudio se corresponde con el enfoque ontosemiótico (EOS) sobre el conocimiento y la instrucción matemática (Godino, 2002; Godino y Batanero, 1994; Godino, Batanero y Font, 2007). Este enfoque parte de la idea de situación-problema que se interpreta, en sentido amplio, como cualquier actividad, pregunta o cuestión que requiera una actividad de matematización. Entonces, se asume que los objetos matemáticos surgen de las prácticas realizadas en la resolución de problemas (Godino y Batanero, 1994).

Para estos autores una práctica matemática es cualquier acción o expresión usada para resolver un problema matemático o generalizarlo o bien comunicar la solución a otros. Además, Godino et al. (2007) diferencian entre prácticas personales o institucionales, es decir, aquellas que son realizadas por un grupo de personas interesadas en resolver un mismo tipo de problemas. Los autores conciben el significado de un objeto matemático como el sistema de prácticas que realiza una persona (significado personal) o se realizan en una institución (significado institucional) para resolver un tipo de situaciones-problema de las que surge el objeto.

En nuestro caso particular, el interés recae sobre el concepto de muestreo y los problemas relacionados con la toma de muestras de una población o la estimación de una característica de la población a partir de las muestras. De estos problemas surgen diferentes objetos matemáticos. Por un lado, el significado institucional que nos interesa es el fijado en los currículos para la educación secundaria, mientras que el significado personal sería el conjunto de prácticas que los estudiantes realizan cuando se enfrentan a situaciones que involucran el concepto de muestreo.

Más concretamente, nos interesamos por determinar cuáles de las prácticas matemáticas realizadas por los estudiantes coinciden con las que se consideran correctas dentro de la institución. En particular, nuestro interés radica en identificar la *comprensión* que una muestra de estudiantes presenta sobre el concepto de muestreo. Entenderemos la comprensión en el sentido de Godino (1996; 2002), es decir, como un rasgo psicológico que no puede ser observado directamente en el sujeto, pero sí puede evaluarse indirectamente a partir de las prácticas personales.

Siguiendo a este autor, así como lo expuesto en el enfoque onto-semiótico (Godino y Batanero, 1994; Godino et al. 2007), la evaluación de la comprensión sería el estudio de la correspondencia entre los significados personales e institucionales de un cierto objeto matemático. Se dirá que el alumno “comprende”, si las prácticas que realiza al proponerle una tarea de muestreo son adecuadas y corresponden con las que se consideran correctas en la institución escolar. Por tanto, siguiendo a estos autores, en este trabajo se evaluará la comprensión sobre el muestreo de los alumnos que participan en el estudio mediante el análisis de las respuestas (entendidas como prácticas), que elaboran al responder a un conjunto de tareas que caracterizan el instrumento de evaluación propuesto para su evaluación.

1.4. OBJETIVOS E HIPÓTESIS DE LA INVESTIGACIÓN

En las líneas anteriores se ha destacado el creciente interés que la enseñanza de la Estadística ha experimentado. En particular, se ha subrayado la importancia de que se comprendan las ideas fundamentales sobre las que se fundamenta. En este sentido, Harradine, Batanero y Rossman (2011) indican que es notable saber relacionar las características de las muestras con las de la población, puesto que es el principal fin de la Estadística. Estos autores sugieren que es posible una comprensión informal del muestreo desde la secundaria, donde los estudiantes debieran comenzar a diferenciar la representatividad y la variabilidad de las muestras, así como el efecto que tiene el tamaño de la muestra sobre la variabilidad de la variabilidad de la misma.

En este sentido, se identifica que el concepto de muestreo adquiere un papel protagonista en el aprendizaje de la Estadística. Por tanto, el *objetivo general* del trabajo es evaluar algunas ideas asociadas al concepto de muestreo en alumnos de Educación Secundaria Obligatoria, en particular. Desde este objetivo general se definen una serie de objetivos específicos que tratan de enmarcar el problema de investigación y guiar en la realización del estudio del mismo.

En primer lugar, la revisión previa de los documentos curriculares conduce a la caracterización del siguiente objetivo:

Objetivo 1. Identificar, describir y comparar los contenidos matemáticos presentados en los documentos curriculares. La finalidad es observar cuáles son los contenidos matemáticos asociados al muestreo que son contemplados por los

documentos curriculares, así como identificar las orientaciones metodológicas consideradas. Este análisis previo guía en la identificación de los contenidos que esperamos que hayan formado parte del proceso de enseñanza y aprendizaje de la etapa de secundaria obligatoria.

Objetivo 2. Realizar una síntesis de las investigaciones previas. La caracterización de este propósito comporta la revisión de la literatura existente. Este análisis se plantea en el Capítulo 2, revelando los resultados que se han obtenido en las investigaciones previas en relación a los contenidos matemáticos que evaluamos. Este estudio nos permite identificar el tipo de tareas que han formado parte de las investigaciones. En este sentido, su consideración propicia la reflexión sobre la selección de las mismas durante el diseño del instrumento de evaluación que aplicamos. Además, esta revisión nos permite valorar si nuestro trabajo puede aportar o complementar nuevos resultados.

Objetivo 3. Evaluar y analizar la comprensión de la relación entre el valor de una proporción en la población y la frecuencia relativa esperada en muestras tomadas de dicha población en alumnos de Educación Secundaria Obligatoria. Nos hemos centrado en la estimación de una proporción, porque es un estadístico muy sencillo y aparece en muchas situaciones comprensibles para el alumnado. Para la evaluación de dichos contenidos hemos considerado la aplicación de un cuestionario, el cual se describe de manera más detallada en el Capítulo 3.

Objetivo 4. Evaluar y analizar la comprensión de la variabilidad en el muestro y el efecto del tamaño de la muestra sobre dicha variabilidad en alumnos de Educación Secundaria Obligatoria. El mismo instrumento de evaluación nos permite evaluar esta comprensión a través del planteamiento de ítems en los que se demanden la generación de muestras pequeñas o grandes. También usaremos fenómenos aleatorios cuyos sucesos elementales sean equiprobables o no equiprobables para comparar los resultados.

Hipótesis iniciales

Aunque nuestro trabajo es de tipo exploratorio, los resultados obtenidos en trabajos realizados con futuros profesores de educación primaria (Gómez, Batanero y Contreras, 2014) nos lleva a plantear las siguientes hipótesis:

H1. Los contenidos curriculares tanto del Decreto de Enseñanzas Mínima como del Currículo básico contemplan o reflejan los aspectos esenciales asociados al concepto de muestreo que son objeto de investigación. Nos apoyamos también en nuestro propio conocimiento del currículo.

H2. Algunos alumnos presentan una comprensión insuficiente de las ideas básicas asociadas al concepto de muestreo. Esperamos que la mayoría de alumnos tenga una buena percepción de la proximidad entre la frecuencia relativa esperada en la muestra y la proporción en la población. Sin embargo, en algunos casos se espera que el sujeto muestre la creencia de que todos los resultados son equiprobables, lo cual guía en la estimación de frecuencias esperadas independientemente de la proporción en que se encuentra dicha frecuencia en la población. Esperamos más dificultad en la idea de variabilidad y es posible que no se perciba el efecto del tamaño de la muestreos sobre la variabilidad. Además, estos resultados se observaron en el estudio de Gómez (2014) con futuros profesores y pensamos se repetirán en el nuestro.

H3. El análisis de las respuestas del instrumento de evaluación refleja un progreso gradual en los contenidos evaluados en los alumnos del 4º curso. La muestra está constituida por alumnos de Educación Secundaria. Además, los individuos que conforman la muestra cursan 2º ESO o 4º ESO. Por tanto, consideramos relevante para la investigación estudiar si existe una diferencia significativa entre los dos grupos que configuran la muestra. Esperamos que por su mayor conocimiento y madurez los alumnos de 4º tengan mejores resultados

1.5. MARCO CURRICULAR

Este apartado está enfocado en realizar una revisión de los documentos curriculares oficiales a nivel estatal, que se concretan a nivel autonómico. Como la parte empírica del trabajo se desarrolla en la comunidad autónoma de Aragón, la revisión de los documentos autonómicos está centrada en los documentos curriculares propios de la comunidad. Además, este análisis permite realizar un estudio complementario donde se observen las discrepancias o puntos en común que se reflejan en ambos currículos.

El foco de interés del análisis de los documentos curriculares se corresponde con la identificación de los contenidos relacionados con el muestreo durante la etapa de la Educación Secundaria Obligatoria (ESO).

Finalmente, señalar que se realiza el estudio a lo largo de la etapa educativa con el objetivo de observar la evolución que experimentan los contenidos correspondientes al bloque de contenidos asociado a este trabajo.

1.5.1. EL MUESTREO EN LOS DECRETOS DE ENSEÑANZA MÍNIMA

Este primer análisis está centrado en el currículo oficial estatal (MEC, 2007) por el que se establece las enseñanzas mínimas correspondientes a la etapa de Educación Secundaria Obligatoria. En primer lugar, el currículo organiza los contenidos de la materia de Matemáticas en bloques de contenidos definidos para cada curso. En concreto, los aspectos asociados a la Estadística quedan descritos bajo el bloque denominado: “Bloque 6. Estadística y probabilidad”. En particular, la información asociada al concepto de muestreo queda recogida en la Tabla 1.1.

Batanero, Gea, Arteaga y Contreras (2014) analizan los contenidos y la metodología sugeridos en las directrices curriculares españolas correspondientes a la enseñanza obligatoria. Este análisis revela que en la Educación Secundaria Obligatoria se sigue contemplando la organización de los datos y gráficos, iniciada en la Educación Primaria. Aunque en el primer curso de la etapa de secundaria se amplían los contenidos, al considerar los conceptos de frecuencia absoluta y relativa, así como nociones de muestreo. Además, identificamos que en este primer curso se incluye el diseño de experiencias para el análisis de conjeturas relacionadas con un fenómeno aleatorio. Por tanto, se puede deducir que la idea que subyace es introducir al alumno el significado frecuencial de la probabilidad adoptando un proceso gradual.

Por otro lado, el currículo estatal contempla explícitamente la consideración de la representatividad en el muestreo desde el tercer curso de la etapa educativa. En relación con la probabilidad, a partir de este curso se incluye como contenido el cálculo de probabilidades mediante la regla de Laplace así como la estimación de las mismas mediante simulación y experimentación. Por tanto, se introduce la probabilidad desde su significado clásico y se refuerza su introducción desde el enfoque frecuencial. En relación con este segundo significado, se identifica que el currículo cita la simulación como técnica para el cálculo de probabilidades.

En el último curso de esta etapa educativa se describen dos vías. Por un lado, se considera que la opción A está enfocada para aquellos alumnos que quieran realizar una

formación profesional. Mientras que la opción B está dirigida a los estudiantes que pretendan continuar con los estudios de Bachillerato. La diferencia radica en una ampliación de los contenidos para esta segunda modalidad citada, los cuales quedan identificados en la Tabla 1.1

Tabla 1.1. Contenidos de muestreo en el Decreto de Enseñanzas Mínimas

Curso	Contenido
1º	<p>Formulación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos y diseño de experiencias para su comprobación.</p> <p>Reconocimiento y valoración de las matemáticas para interpretar y describir situaciones inciertas.</p> <p>Diferentes formas de recogida de información. Organización en tablas de datos recogidos en una experiencia. Frecuencias absolutas y relativas (MEC, 2007, p.127).</p>
2º	<p>Diferentes formas de recogida de información. Organización de los datos en tablas. Frecuencias absolutas y relativas, ordinarias y acumuladas.</p> <p>Medidas de centralización: media, mediana y moda. Significado, estimación y cálculo. Utilización de las propiedades de la media para resolver problemas. Utilización de la media, la mediana y la moda para realizar comparaciones y valoraciones (MEC, 2007, p.129).</p>
3º	<p>Necesidad, conveniencia y representatividad de una muestra. Métodos de selección aleatoria y aplicaciones en situaciones reales. Atributos y variables discretas y continuas.</p> <p>Agrupación de datos en intervalos. Histogramas y polígonos de frecuencias.</p> <p>Construcción de la gráfica adecuada a la naturaleza de los datos y al objetivo deseado.</p> <p>Media, moda, cuartiles y mediana. Significado, cálculo y aplicaciones.</p> <p>Análisis de la dispersión: rango y desviación típica. Interpretación conjunta de la media y la desviación típica.</p> <p>Utilización de las medidas de centralización y dispersión para realizar comparaciones y valoraciones. Actitud crítica ante la información de índole estadística.</p> <p>Experiencias aleatorias. Sucesos y espacio muestral. Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar.</p> <p>Cálculo de la probabilidad mediante la simulación o experimentación.</p> <p>Utilización de la probabilidad para tomar decisiones fundamentadas en diferentes contextos. Reconocimiento y valoración de las matemáticas para interpretar, describir y predecir situaciones inciertas (MEC, 2007, p.132).</p>
4º A	<p>Análisis elemental de la representatividad de las muestras estadísticas.</p> <p>Utilización de las medidas de centralización y dispersión para realizar comparaciones y valoraciones.</p> <p>Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar (MEC, 2007, p.135).</p>
4º B (se añade)	<p>Análisis crítico de tablas y gráficas estadísticas en los medios de comunicación. Detección de falacias.</p> <p>Representatividad de una distribución por su media y desviación típica o por otras medidas ante la presencia de descentralizaciones, asimetrías y valores atípicos. Valoración de la mejor representatividad en función de la existencia o no de valores atípicos (MEC, 2007, p.137).</p>

En relación con lo recogido en el Boletín Oficial de Aragón (BOA) elaborado desde el Departamento de Educación, Cultura y Deporte (2007), se contempla cada uno de los contenidos señalados anteriormente (ver Tabla 1.1.). El análisis revela que el currículo autonómico amplía dichos contenidos. Por tanto, la Tabla 1.2. muestra aquellos puntos que son añadidos.

Tabla 1.2. Contenidos relacionados con el muestreo en el Decreto de Enseñanzas Mínimas de la Comunidad de Aragón

Curso	Contenido
1º	Población y muestra. Características cualitativas y cuantitativas de una población. Distribuciones discretas (Departamento de Educación, Cultura y Deporte, 2007, p.237).
3º	Ánálisis de las características de una población a partir de tablas y gráficos estadísticos. Imprevisibilidad y regularidad. Frecuencia relativa y probabilidad de un suceso: estabilidad de las frecuencias (Departamento de Educación, Cultura y Deporte, 2007, pp.247-248).
4ºB	Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con estudios estadísticos de poblaciones y con el azar (Departamento de Educación, Cultura y Deporte, 2007, pp.257-258).

En términos genéricos, la lectura de las directrices recogidas en el currículo autonómico refleja un mayor nivel de especificidad en relación a los contenidos y destrezas que tienen que trabajarse bajo ese bloque de contenidos a lo largo de la etapa educativa.

En relación con la información resumida en la Tabla 1.2., identificamos la inclusión de contenidos relacionados con la idea de muestreo. Por un lado, se espera que el alumno de tercer curso detecte la imprevisibilidad y regularidad, que constituyen una característica inherente a la probabilidad enfocada desde el significado frecuencial. Por otro lado, el currículo autonómico, en el último curso, considera la utilización de un lenguaje adecuado para describir y cuantificar a informaciones asociadas a un estudio estadístico.

Orientaciones metodológicas

Las directrices curriculares a nivel estatal y autonómico remarcan que la introducción de los contenidos correspondientes al bloque de estadística y probabilidad debe seguir un proceso gradual. Entonces, los primeros cursos suponen una aproximación al estudio de fenómenos aleatorios sencillos mediante la experimentación y el tratamiento, por medio de tablas y gráficas, de datos estadísticos. Las orientaciones

correspondientes a la comunidad de Aragón resaltan que el objetivo principal de la enseñanza de la estadística descriptiva no es el desarrollo de destrezas de cálculo o de la habilidad para elaborar gráficos estadísticos, sino que está enfocada en el análisis cualitativo de las características de la población y el fomento de la actitud crítica. Por tanto, se concede mayor importancia a la interpretación de los datos y los resultados que a la memorización y el cálculo.

El currículo a nivel autonómico sugiere que los datos se obtengan del entorno del alumno y contempla algunas de las ideas fundamentales descritas por Burril y Biehler (2011). En concreto, se expresa la importancia que adquieren los datos en el estudio estadístico, subrayando la importancia del contexto (Moore, 1991). Las directrices curriculares autonómicas indican la necesidad de introducir al alumno conceptos matemáticos asociados a la inferencia estadística desde un enfoque informal, basado en la simulación:

Además del enfoque descriptivo, en la estadística también tiene cabida el enfoque inductivo, que tiene por objetivo el conocimiento de las características de una población a partir de las observaciones realizadas sobre muestras de ella. El instrumento para este punto de vista de la estadística es la teoría de probabilidades y el estudio de las distribuciones teóricas de probabilidad, y queda fuera del alcance del nivel de la Educación secundaria obligatoria. No obstante, puede hacerse un tratamiento informal de la relación entre características de una población y de una muestra y una aproximación experimental a la idea de representatividad" (Departamento de Educación, Cultura y Deporte, 2007, p.264).

Estas líneas reflejan la conveniencia de guiar al alumno hacia la inferencia estadística aunque no se trabaje de manera exhaustiva. Como se indica en el fragmento, la finalidad es que el alumno reflexione sobre la idea de representatividad de una muestra en relación a la población de la que se obtiene. Esta idea junto con la de variabilidad de la muestra supone dos ideas relevantes del muestreo, las cuales son objeto de evaluación del presente trabajo. De hecho, Batanero et al. (1994) indican que una comprensión adecuada del muestreo implica el equilibrio adecuado de los dos términos descritos anteriormente.

1.5.2. EL MUESTREO EN EL CURRÍCULO BÁSICO

Bajo este apartado se realiza un análisis similar al caso anterior con respecto al Currículo Básico (MEC, 2015). Análogamente, se exponen los contenidos de muestreo descritos por el currículo para el "Bloque5. Estadística y probabilidad". Por otro lado, tras la presentación de los contenidos recogidos en el bloque para cada curso, se señalan

los contenidos, que por primera vez, se incluyen en el currículo oficial en comparación con el documento curricular anteriormente presentado (MEC, 2007).

En primer lugar, una de las novedades que introduce la Ley Orgánica para la Mejora de la Calidad Educativa (LOMCE) son las dos vías que se inician en el tercer curso: “matemáticas orientadas a las enseñanzas académicas” y “matemáticas orientadas a las enseñanzas aplicadas”. Esta diferenciación radica en que el enfoque con el cual se presentan las segundas tiene que responder a un interés por mostrar la aplicación práctica de los contenidos matemáticos en situaciones reales frente a la profundización en los aspectos teóricos (MEC, 2015, p.339).

Tabla 1.3. Contenidos relacionados con el muestreo en el Currículo Básico

Curso	Contenido
1º y 2º ESO	<p>Población e individuo. Muestra. Variables estadísticas. Variables cualitativas y cuantitativas. Frecuencias absolutas y relativas. (No incluye las frecuencias acumuladas descritas en 2ºESO).</p> <p>Medidas de tendencia central. Medidas de dispersión.</p> <p>Fenómenos deterministas y aleatorios. Formulación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos y diseño de experiencias para su comprobación. Espacio muestral en experimentos sencillos.</p> <p>Frecuencia relativa de un suceso y su aproximación a la probabilidad mediante la simulación o experimentación. Sucesos elementales equiprobables y no equiprobables (MEC, 2015, p.413).</p>
3º curso Matemáticas orientadas a las enseñanzas académicas	<p>Fases y tareas de un estudio estadístico. Población, muestra. Variables estadísticas: cualitativas, discretas y continuas.</p> <p>Métodos de selección de una muestra estadística. Representatividad de una muestra. Frecuencias absolutas, relativas y acumuladas. Parámetros de posición. Cálculo, interpretación y propiedades. Parámetros de dispersión.</p> <p>Experiencias aleatorias. Sucesos y espacio muestral.</p> <p>Utilización de la probabilidad para tomar decisiones fundamentadas en diferentes contextos (MEC, 2015, p.394).</p>
3º curso Matemáticas orientadas a las enseñanzas aplicadas	<p>Fases y tareas de un estudio estadístico. Población, muestra. Variables estadísticas: cualitativas, discretas y continuas.</p> <p>Métodos de selección de una muestra estadística. Representatividad de una muestra. Frecuencias absolutas, relativas y acumuladas. Parámetros de posición: media, moda, mediana y cuartiles. Cálculo, interpretación y propiedades.</p> <p>Parámetros de dispersión: rango, recorrido intercuartílico y desviación típica. Cálculo e interpretación. Interpretación conjunta de la media y la desviación típica (MEC, 2015, p.403).</p>
4º curso Matemáticas orientadas a las enseñanzas académicas	<p>Probabilidad simple y compuesta. Sucesos dependientes e independientes.</p> <p>Experiencias aleatorias compuestas. Probabilidad condicionada.</p> <p>Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar y la estadística.</p> <p>Ánalisis crítico de tablas y gráficas estadísticas en los medios de comunicación. Detección de falacias.</p> <p>Medidas de centralización y dispersión: interpretación, análisis y utilización. Comparación de distribuciones mediante el uso conjunto de medidas de posición y dispersión (MEC, 2015, p.398).</p>
4º curso	Interpretación, análisis y utilidad de las medidas de centralización y dispersión.

Matemáticas orientadas a las enseñanzas aplicadas	Comparación de distribuciones mediante el uso conjunto de medidas de posición y dispersión. Azar y probabilidad. Frecuencia de un suceso aleatorio. Probabilidad simple y compuesta. Sucesos dependientes e independientes (MEC, 2015, p.407).
----------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

En comparación con el currículo oficial de la Ley Orgánica de Educación (LOE), se incluye la simulación desde el primer curso de esta etapa educativa. Además, el último curso contempla la introducción a la correlación independientemente de la modalidad.

Por otro lado, la introducción de la probabilidad desde el significado clásico se adelanta, iniciándose en los primeros cursos de la ESO. En particular, el currículo indica que el cálculo de probabilidades se apoya en la aplicación de la Regla de Laplace (MEC, 2015). De modo que el alumno que continúe la vía orientada a las enseñanzas académicas aprenda desde el tercer curso conceptos matemáticos del análisis combinatorio.

1.5.3. CONCLUSIONES

En términos generales, se observa que la organización de los contenidos asociados al bloque de estadística y probabilidad de la ESO muestra una introducción gradual de los mismos. Un aspecto a destacar es la importancia que se concede a que el alumno desarrolle una actitud crítica ante la información estadística. De modo que el alumno sea capaz de interpretar y evaluar la información estadística que recibe en distintas situaciones, como son los medios de comunicación. Por otro lado, el currículo concede también importancia al lenguaje, siendo el alumno capaz de comunicarse. Estos dos aspectos guardan similitud con las dos competencias que, según Gal (2002), caracterizan la cultura estadística.

El análisis comparativo de los dos documentos curriculares pone en relieve que las directrices del nuevo currículo estatal se caracterizan por considerar un conjunto de contenidos más amplios en cada curso o anticipar la enseñanza de determinados contenidos, siendo este hecho más acusado o relevante en la modalidad definida como “matemáticas orientadas a las enseñanzas académicas” (MEC, 2015).

1.6. CONTENIDOS SOBRE MUESTREO QUE SE CONSIDERAN EN EL TRABAJO

Como se ha indicado, el concepto de muestreo nos introduce en la inferencia y supone un puente que conecta la Estadística y la Probabilidad. Según Batanero (2004), el conocimiento científico tiene una naturaleza pragmática, entendiendo que su construcción se debe a las experiencias empíricas, las cuales suelen ser limitadas. Este hecho implica que las conclusiones deben aportar una información más amplia que los datos utilizados u obtenidos a partir de las observaciones. En el caso particular de la Estadística, el proceso de muestreo proporciona información sobre la población, y por tanto, aumenta nuestro conocimiento sobre la misma, siendo éste el objetivo principal del trabajo estadístico.

Como se ha señalado, Batanero et al. (1994) indican que la aprehensión del muestreo implica el equilibrio adecuado entre dos ideas que aparentemente parecen opuestas: la representatividad muestral y la variabilidad muestral. La primera idea permite señalar que la muestra si es representativa tendrá características similares a las de la población. Esto exige que se haya elegido a través de técnicas adecuadas. Por otro lado, la variabilidad muestral implica que no todas las muestras son iguales. Siguiendo a estos autores, el punto adecuado de equilibrio entre los extremos de información total e información nula en relación a una población es complejo, dependiendo de tres factores: variabilidad de la población, tamaño de la muestra y coeficiente de confianza.

Los ítems que consideramos se centran en la proporción de valores en una muestra que cumplen una cierta condición. Por tanto, intuitivamente se está trabajando con una población con sólo dos resultados: éxito y fracaso. Nos interesamos por la estimación que hacen los estudiantes del valor esperado de éxitos en una muestra de tamaño n , cuando se conoce la proporción de éxitos p en la población y también, cómo estiman este valor p en el caso de que conozca la composición de una muestra de tamaño n .

Por último estamos interesados en analizar si los estudiantes comprenden la variabilidad del muestreo y cómo la reproducen cuando se les pide dar la composición de cuatro muestras del mismo tamaño n en la población de partida.

CAPÍTULO 2.

ANTECEDENTES

2.1. INTRODUCCIÓN

Este capítulo supone una síntesis de las investigaciones más relevantes sobre la comprensión del muestreo, así como de otros conceptos que exigen de una comprensión adecuada del primero, como es el caso del concepto de distribución, y en particular, del concepto de distribución muestral. Dichas investigaciones nos servirán de apoyo para analizar los resultados de nuestro estudio. Con el propósito de realizar esta síntesis, hemos partido de las investigaciones recogidas en los trabajos de Harradine et al. (2011) y Ben-Zvi, Bakker y Makar (2015), principalmente.

En primer lugar, se analizan los niveles de generalidad de un mismo concepto que caracterizan la inferencia estadística. Este análisis revela la importancia y complejidad que adquiere el concepto de distribución y las posibles dificultades que pueden tener los estudiantes al pasar de un nivel a otro del término.

En segundo lugar, el análisis se centra en la revisión bibliográfica de aquellas investigaciones centradas en los conceptos de muestra y muestreo. Este análisis revela las concepciones erróneas que los individuos tienen sobre algunas propiedades de la muestra y el muestreo. Además, se citan algunas de las investigaciones más relevantes llevadas a cabo con estudiantes, donde el interés recae en analizar la comprensión del concepto de muestra al llevar a cabo un muestreo apoyado en la simulación con herramientas de software específico, fundamentalmente.

La reflexión llevada a cabo en el segundo apartado conduce a la identificación de la complejidad del término distribución muestral sobre el que se centra el aprendizaje de la inferencia estadística. Por consiguiente, el siguiente punto se centra en presentar algunas de las investigaciones cuyo objetivo supone el análisis de la comprensión de dicho concepto por los estudiantes.

Por otra parte, la consideración del enfoque frecuencial de la probabilidad pone en relieve la estrecha relación que se establece entre la probabilidad y la estadística. De hecho, se identifica dicho vínculo cuando se pide estimar la proporción en una población a partir de la muestra. Recíprocamente, estimar una probabilidad teórica a partir de la frecuencia relativa en una muestra grande de ensayos es un problema de

muestreo e inferencia. Por esta razón, la última de las revisiones bibliográficas se centra en mostrar aquellas investigaciones enfocadas en la comprensión de la probabilidad desde su enfoque frecuencial.

Finalmente, se sintetizan aquellos aspectos más relevantes que se desprenden de la literatura identificada, señalando los puntos que son considerados para la elaboración del instrumento de evaluación y para el análisis de las respuestas de los alumnos.

2.2. NIVELES DE USO DE LOS CONCEPTOS DE INFERENCIA

La inferencia estadística se fundamenta en la aplicación de dos tipos de procedimientos o técnicas, principalmente, en función de la finalidad u objetivo del problema. Por un lado, si el objetivo es proporcionar un margen de variación para un determinado parámetro desconocido de la distribución, entonces la finalidad es determinar un intervalo de confianza. Por otro lado, el objetivo puede ser confirmar o rechazar una determinada afirmación acerca de la distribución de probabilidad del fenómeno estudiado. Entonces, el interés implica la realización de un contraste de hipótesis.

Un primer análisis de las caracterizaciones citadas pone en relieve los distintos conceptos que subyacen al quehacer del estadístico. Como señalan Liu y Thompson (2009), estos procedimientos exigen un esquema coordinado entre diversos conceptos: la probabilidad, la variabilidad aleatoria, la distribución de los valores de un estadístico muestral, las hipótesis nula y alternativa, el p-valor, el nivel de significación (o bien el coeficiente de confianza y el intervalo), y la lógica de la inferencia.

Harradine et al. (2011) identifican que la inferencia estadística está constituida por tres elementos fundamentales que interactúan entre ellos, los cuales son: el proceso de razonamiento, los conceptos y los cálculos asociados. En relación con los conceptos, estos autores subrayan los conceptos tales como, la distribución, la media, la varianza, la incertidumbre, la aleatoriedad y el muestreo, además de otros específicos del intervalo de confianza o el contraste de hipótesis. Además, indican que el proceso de cálculo se encuentra actualmente resuelto por la tecnología. Por ello, y siguiendo a Garfield y Gal (1999), pensamos que la enseñanza debe favorecer que el individuo sea capaz tanto de utilizar las ideas fundamentales de estadística como de desarrollar un razonamiento estadístico adecuado con el fin de dar sentido a la información estadística.

Todos los autores mencionados coinciden en que la finalidad de la inferencia estadística es obtener información de la población a través de las muestras. Esta finalidad exige la reflexión sobre el conjunto de datos que se van a recoger con el fin de elaborar conclusiones con un grado de probabilidad adecuado (generalmente, una probabilidad condicional a un modelo). Para conseguir este fin, es necesario trabajar con tres tipos de distribuciones asociadas a la inferencia estadística que están presentes durante todo el proceso (Harradine et al., 2011; Kadijevich, Kokol-Voljc y Lavicza, 2008):

- La *distribución teórica de probabilidad* que modela los valores de una variable aleatoria tomada de una población o un proceso. Dicha distribución depende de algún valor del *parámetro* que es generalmente desconocido. Con frecuencia, los problemas estadísticos relativos a una población son descritos por una variable aleatoria unidimensional con distribución teórica de tipo normal. En este caso, la distribución queda especificada por dos parámetros, su media y desviación típica, denotados por μ, σ , respectivamente. En otros casos, se trabajará con la distribución binomial, uniforme o exponencial, entre otras. En nuestro caso, se considera la Binomial, cuyo principal parámetro es la proporción de éxitos p .
- La *distribución del conjunto de datos* que se obtiene a partir de la realización de varias repeticiones independientes del experimento aleatorio. Este conjunto de datos recibe el nombre de *muestra aleatoria simple*. Desde esta muestra, se definen distintos estadísticos tales como la media muestral y la desviación típica muestral, que se denotan \bar{x} y s , respectivamente. Estos estadísticos pueden aplicarse en el procedimiento para determinar los valores aproximados que toman los parámetros de la distribución teórica o bien para contrastar hipótesis sobre dichos parámetros. En los problemas planteados en nuestro estudio se pide a los alumnos pensar en cuatro muestras del mismo tamaño. En este caso, el estadístico correspondiente es la proporción de éxitos en la muestra (nosotros trabajamos con una forma equivalente que es el número de éxitos en cada muestra).
- La *distribución en el muestreo* de un estadístico se define como la distribución de probabilidad de todos los posibles valores que puede tomar el estadístico muestral en relación con las posibles muestras que constituyen el espacio muestral asociado al proceso o población que se desea estudiar. Aunque para una muestra particular tenemos un solo valor del estadístico, al pensar en las infinitas muestras que se

pueden tomar de una población, tenemos una variable aleatoria; la distribución de esta variable aleatoria es la distribución en el muestro. Por ejemplo, si la muestra de tamaño n se tomó de la distribución normal $N(\mu, \sigma)$, la distribución en el muestreo de la media de la muestra \bar{x} tiene una distribución normal $N(\mu, \frac{\sigma}{\sqrt{n}})$. En nuestro caso, sólo se pide cuatro muestras, cuyo número de éxitos son elementos de la distribución muestral.

La necesidad de trabajar en el mismo problema con las tres distribuciones anteriores, para las que se define una media y una desviación típica, origina una dificultad relevante para los estudiantes. Por otro lado, Batanero et al. (1994) señalan que en estadística descriptiva la unidad de análisis es cada dato (referido a una persona u objeto). La media de una muestra es la media de los datos que constituyen la muestra. No obstante, al comenzar el estudio de la inferencia, el interés incide en conocer información sobre la población y determinar la media teórica de la población. Este nuevo enfoque exige la obtención de distintas muestras de una población particular. En este caso, la unidad de análisis es cada una de las posibles muestras de un mismo tamaño, que se toman desde la población. Entonces, la media de cada muestra conduce al estudio de la distribución en el muestreo del estadístico denominado media muestral. De modo que, el grado de complejidad asociado a este último análisis se corresponde con un nivel de abstracción mayor. De hecho, los estudiantes confunden con frecuencia estos diferentes niveles de concreción del concepto de distribución en estadística descriptiva e inferencia, lo que es un problema relacionado con la enseñanza del muestreo (Schuyten, 1991).

Batanero, Díaz, Contreras y Roa (2013) al igual que Harradine et al. (2011) indican que el razonamiento inferencial requiere la coordinación de distintos conceptos: los datos, la población de donde se tomaron y las posibles muestras de la misma. Por tanto, exige conectar las tres distribuciones citadas anteriormente. A su vez, la idea de muestra y distribución muestral requieren de una adecuada comprensión del concepto de muestreo. Como señala Wild y Pfannkuch (1999), uno de los modos fundamentales del razonamiento estadístico implica la percepción de la variabilidad debida a diversas fuentes como pueden ser las muestras seleccionadas mediante el proceso de muestreo. Las ideas fundamentales (Burril y Biehler, 2011) de distribución y la variabilidad son la base de la comprensión de la variabilidad y representatividad muestral que sostiene la

consideración de la distribución muestral como objeto de análisis para inferir información de la distribución teórica (Ben-Zvi et al., 2015).

2.3. COMPRENSIÓN DE LAS PROPIEDADES DEL MUESTREO

Desde el campo de la psicología se han llevado a cabo investigaciones sobre el análisis de los errores referidos al razonamiento inferencial centrado en el contexto de la toma de decisiones en situaciones de incertidumbre (Harradine et al., 2011). Un resumen de estos estudios se presenta en Kahneman, Slovic, y Tversky (1982). Los errores se corresponden con el empleo de *heurísticas* en la resolución de los problemas. En esta línea de investigación se concibe el término heurística como una regla de conducta, de carácter generalmente inconsciente, que guía la resolución de las tareas de razonamiento complejas, porque reducen la información de las mismas. Por tanto, estos procesos cognitivos reducen el nivel de complejidad de la tarea durante el proceso de resolución de la misma.

En particular, se describen tres heurísticas fundamentales: la representatividad, la disponibilidad y “ajuste y anclaje”. Además, se estudian los sesgos asociados, así como sus implicaciones teóricas y prácticas. Las siguientes líneas describen las dos primeras heurísticas citadas.

- La *heurística de representatividad* (Tversky y Kahneman, 1982) implica que la probabilidad de obtener una muestra se estima comparando la similitud que guarda con la población a la que pertenece. Los sesgos asociados son la *insensibilidad al tamaño de la muestra* y la *falacia del jugador* (Tversky y Kahneman, 1974). El primer sesgo implica que la persona asume que una muestra, aunque sea pequeña, siempre representa a la población, lo cual conduce a graves consecuencias en el trabajo estadístico. En este sentido, Kahneman et al. (1982) expresan que la persona cree en la “ley de los pequeños números”, lo cual conlleva a la consideración de intervalos de confianza con una amplitud demasiada baja. Además, esperan que los primeros resultados sean similares a los futuros. Por otro lado, la *falacia del jugador* supone considerar que el resultado de un experimento aleatorio afectará en la probabilidad de sucesos futuros.
- La *heurística de la disponibilidad* (Tversky y Kahneman, 1974) consiste en estimar la verosimilitud de un suceso basándose en la información previa de la persona, es decir, los recuerdos. La aplicación de dicha heurística supone la estimación de la probabilidad en función de la facilidad por encontrar ejemplos de situaciones

similares (Díaz, 2003). Se identifican tres sesgos consecuencia de la aplicación de esta creencia. En primer lugar, *el sesgo de equiprobabilidad* (Lecoutre, 1992) que consiste en pensar que los resultados de cualquier fenómeno aleatorio son igualmente probables. Otro de los sesgos es el *enfoque del resultado* (Konold, 1989) que implica interpretar el resultado de un experimento desde el enfoque subjetivo de la probabilidad. En otros términos, el sujeto no interpreta la situación desde un enfoque frecuencial, por lo que su respuesta supone una predicción de la misma en función del grado de creencia que otorga a la obtención de dicho suceso. Finalmente, la *ilusión de control* (Langer, 1982) se corresponde con el hecho de que el sujeto no distingue la diferencia entre los juegos de habilidad y los de azar. Por tanto, el sujeto tiene la creencia de poder controlar el azar.

2.4. COMPRENSIÓN DEL MUESTREO

Tras la identificación de los principales errores referidos al razonamiento inferencial en términos de heurísticas, las siguientes líneas suponen una síntesis de algunas de las investigaciones centradas en observar cómo los estudiantes interpretan las ideas de muestra y muestreo.

En primer lugar, la investigación de Watson y Moritz (2000) analiza la idea intuitiva que una muestra formada por un conjunto de 62 alumnos (8-9 años, 11-12 años, 14-15 años) manifiesta sobre el concepto de muestra. Se identifica que los estudiantes tienen dificultad al realizar la transición hacia el significado que adquiere el término de muestra en estadística. Aunque pueden hacer generalizaciones adecuadas desde una muestra pequeña en relación a la población a la que pertenecen, esta idea intuitiva no incluye aceptar la variabilidad intrínseca al concepto de muestra (Ben-Zvi et al., 2015).

Desde esta observación, Watson y Moritz (2000) indican la necesidad de hacer explícitas las diferencias del concepto en función del contexto en el que aparece, como por ejemplo, comparar las diferencias entre tomar una muestra pequeña de comida la cual está asociada a una población homogénea, con la obtención de una muestra de una población de sujetos para analizar una característica de la misma, donde las muestras tomadas presentan una variabilidad mayor. En otro estudio, Watson (2004) expone que los estudiantes conceden mayor legitimidad a las muestras tomadas con algún

procedimiento sesgado, como por ejemplo, muestras voluntarias que a aquellos métodos que generan muestras aleatorias, como el muestreo aleatorio estratificado.

Shaughnessy, Ciancetta y Canada (2004) realizan una investigación sobre la comprensión que los estudiantes tienen sobre el concepto de variabilidad en estadística. La muestra está constituida por 272 estudiantes de entre 10-19 años. Los resultados se obtienen desde el análisis de una serie de tareas sobre muestreo combinado con entrevistas personales. El análisis de los resultados da lugar a la identificación de tres tipos de razonamientos denominados por los autores aditivo, proporcional y distribucional. Como resultado se observa que la mayoría de los alumnos se sitúan en el primero de los razonamientos citados, que se relaciona con aquellas respuestas que se guían por la proporción en la que se presenta en la población. En relación con el concepto de variabilidad, se identifica que muchos de los estudiantes creen en la posibilidad de resultados muy variados independiente del tamaño de la muestra. En contraposición, otros otorgan muy poca variabilidad al conjunto de posibles resultados que se puede obtener en un muestreo.

2.5. COMPRENSIÓN DE LA DISTRIBUCIÓN MUESTRAL

Como se ha citado anteriormente, el concepto de distribución muestral es un tópico abstracto, que requiere que el estudiante coordine distintos conceptos como: muestra, distribución, variabilidad y muestreo. Esta dificultad conduce a que muchos estudiantes no presenten una comprensión suficiente del mismo o no sean capaces de interpretar los resultados de los cálculos realizados.

La investigación realizada por Saldanha y Thompson (2002) tiene como finalidad desarrollar un análisis epistemológico de las ideas de muestra, distribución muestral y márgenes de error a partir de una muestra de alumnos de 16 a 18 años. Para lo cual, los autores llevan a cabo una experiencia de enseñanza en la que usan software que permite la generación de muestras aleatorias de una población. De este modo los alumnos pueden generar un gran número de muestras y estudiar sus características. La investigación se centra en analizar el modo en que los alumnos coordinan los tres niveles de la idea de distribución (distribución de la población, distribución de la muestra y distribución del estadístico en el muestreo).

Los autores identifican que algunos estudiantes tienen una imagen aditiva de la muestra, porque tienden a concebir la muestra como un conjunto de una población y múltiples muestras como múltiples subconjuntos disjuntos. Mientras que algunos alumnos presentan una concepción multiplicativa de la muestra (MCS), en la que se relaciona el concepto de muestra con la población, así como con la distribución muestral. En contraste con el primero, la MCS conlleva una relación de imágenes más rica que favorece la aceptación de muestras que sean relativamente inusuales; por tanto, ayuda a la comprensión de la inferencia estadística. No obstante, se detecta que la mayoría de los estudiantes se orientan hacia una imagen aditiva del concepto de muestra.

En definitiva, Saldanha y Thompson (2002) identificaron que los estudiantes no presentaban una distinción clara entre los distintos niveles asociados al concepto de distribución, anteriormente descritos. Por tanto, estos autores exponen que la utilización de software no es suficiente para conseguir una adecuada comprensión del concepto. Por otro lado, recomiendan que los estudiantes se familiaricen con el proceso de muestreo, además de permitirles que exploren los conceptos mediante actividades no guiadas. De este modo, el alumno puede construir y evaluar su propio conocimiento. Otro aspecto relevante es proporcionar un espacio en el que los estudiantes puedan discutir las observaciones una vez que la tarea ha concluido.

Chance, delMas y Garfield (2004) presentan un proyecto de investigación que tiene como finalidad analizar si existe una mejora en el razonamiento sobre la distribución muestral en un grupo estudiantes que van a iniciar los estudios superiores, a lo largo de un curso en que se usan herramientas de software específica. Los autores identifican que los estudiantes no son capaces de explicar los conceptos que están en juego en los procedimientos que aplican, a pesar de que pueden aprender como calcular un intervalo de confianza y lleven a cabo test de hipótesis. Es decir, aplican las técnicas como si fueran una receta, pero no saben describir el por qué de tales procedimientos.

En este sentido, los autores identifican que una de las razones de esta dificultad se debe a que los alumnos no comprenden los conceptos de distribución muestral y variabilidad de las muestras, entre otros. De acuerdo con estos autores, para lograr una comprensión suficiente de la idea de distribución muestral es necesario reflexionar previamente sobre los conceptos sobre los que se basa el término, con la finalidad de

interpretarlos de manera correcta en el contexto de las distribuciones muestrales. Los conceptos previos que señalan son: la muestra, el estadístico en la muestra, la variabilidad de las muestras y el efecto del tamaño sobre la variabilidad. Con respecto a este último apunte, citar que algunos de estos conceptos son objeto de evaluación en este trabajo. Como consecuencia de esta investigación, se observa también que los estudiantes no identifican las diferencias entre los distintos niveles conceptuales asociados al concepto de distribución, este resultado se identifica también en la investigación de Saldanha y Thompson (2002).

Bakker (2004) lleva a cabo una investigación centrada en el análisis de un proceso de instrucción. El objetivo es que los sujetos (alumnos de 13-14 años) desarrollen la noción de distribución en el muestreo mediante el razonamiento fundamentado de las gráficas asociadas a las muestras obtenidas, con la particularidad de que el tamaño de las muestras va aumentando. El autor concluye que la enseñanza favoreció que los alumnos comprendieran que un mayor tamaño de la muestra implica una menor variabilidad y, por tanto, la muestra representa mejor a la población. Según Bakker (2004), permitir al alumno la formulación de conjeturas sobre las muestras, fundamentada en el ciclo de investigación (Wild y Pfannkuch, 1999), estimula que el alumno ponga en marcha el esquema conceptual para predecir la distribución, lo cual beneficia el desarrollo del razonamiento sobre las muestras.

Más recientemente, Vanhoof, Castro Sotos, Onghena y Verschaffel (2007) se plantearon analizar la efectividad de un software denominado *Sampling Activity Distribution* (SDA), desarrollado por Chance et al. (2004) y delMas, Garfield y Chance (2004). La muestra estuvo constituida por 221 estudiantes universitarios que estaban cursando el segundo curso de la carrera. Además, el plan de estudios incluye tres asignaturas en las que se imparten contenidos de estadística.

La comparación de los resultados del pre-test y pos-test revela una mejora significativa de la comprensión de la distribución muestral. En relación con la noción de distribución muestral, se identifica que solamente la mitad de los sujetos muestran una comprensión adecuada. No obstante, los autores señalan que esta mejora puede deberse a otros aspectos o variables que no se controlan en el estudio, como por ejemplo el campo de estudio de los sujetos, así como su experiencia previa, puesto que habían sido instruidos en conceptos propios de estadística.

2.6. COMPRENSIÓN DE LA PROBABILIDAD, DESDE EL PUNTO DE VISTA FRECUENCIAL

Siguiendo a Ben-Zvi et al. (2015), la idea de variabilidad que subyace en los conceptos de muestra y muestreo se fundamenta en la aleatoriedad y el azar, o el estudio de la probabilidad. Por otro lado, la Ley de los Grandes Números garantiza que las muestras con un tamaño mayor representan mejor a la población de la que fue tomada. De modo que los estadísticos asociados a la muestra están más próximos a los valores de los parámetros de la población. Entonces, la Ley de los Grandes Números supone la base del significado de la probabilidad desde el enfoque frecuencial, así como se encuentra implícita en algunas de las ideas asociadas al concepto de muestreo. De este modo, se justifica el interés por realizar una revisión sobre aquellas investigaciones centradas en el análisis de la comprensión de la probabilidad desde el enfoque frecuencial.

Desde la revisión de las mismas, se identifica que muchas de las tareas usadas para la investigación sobre la comprensión de la probabilidad desde el significado frecuencial pueden reinterpretarse en término de tareas de evaluación de la comprensión de ideas básicas de muestreo. De hecho, los ítems que constituyen nuestro cuestionario están adaptados de algunas de estas investigaciones.

En primer lugar, citamos la investigación realizada por Green (1983a), fundamentada en la elaboración de un cuestionario dirigido a niños ingleses (de 11 a 16 años) donde se incluye un ítem sobre la comprensión de la estimación frecuencial de la probabilidad asociado un experimento con sucesos no equiprobables (lanzar al aire 100 chinchetas para determinar cuántas caen con la punta hacia arriba o hacia abajo).

En la prueba piloto (Green, 1983b) preguntó a 66 niños que estimaran, el número de chinchetas que esperan que caigan con la punta hacia arriba y cuántas con la punta hacia abajo cuando se lanzan 100 chinchetas sobre una mesa. Además, se les proporciona información previa, puesto que se les proporciona los datos asociados a la realización del experimento (donde 68 de 100 chinchetas caen con la punta hacia arriba). Por tanto, se espera que los niños den un resultado parecido, pero no idéntico. No obstante, los resultados revelaron que el 61% de los estudiantes mostraron el sesgo de equiprobabilidad, es decir, proporcionaron respuestas en las que la mitad de las chinchetas, aproximadamente, caerían hacia arriba, sin tener en cuenta la información

frecuencial. Por otro lado, el 15% de los alumnos mostraron una preferencia a la punta hacia arriba (correcta o parcialmente correcta), pero dando una cantidad muy alejada de la esperada. Tras los diferentes estudios pilotos y entrevistas con los niños, el ítem se cambió un poco. A pesar de estas modificaciones, no existe una mejora del sesgo citado, el cual se mantenía en todas las edades.

La investigación de Green (1983a) constituida por una muestra de 2930 alumnos, que supone una continuación de la citada anteriormente, refleja que el 64% presenta el sesgo de equiprobabilidad, y solamente el 17% dan una estimación correcta de tipo frecuencial. Luego, se observa que las tasas de acierto y sesgo no mejoraron con esta versión definitiva.

Por otro lado, la investigación realizada por Cañizares (1997) sobre una muestra de 253 niños españoles (11-14 años) considera el ítem propuesto por Green (1983a). En este caso, los resultados reflejan que el 64.1% de los alumnos muestran el sesgo de equiprobabilidad, mientras que sólo el 15% responde correctamente. Además, se observa un aumento no significativo con la edad en la consideración del principio de equiprobabilidad. En relación con este resultado, se justifica si se contempla el hecho que en aquella época no se iniciaba el estudio de la probabilidad hasta los 14 años, por lo que la mayoría de los niños nunca la había estudiado. No obstante, la revisión del currículo oficial, expuesta en el Capítulo 1, revela que actualmente algunos contenidos de probabilidad son contemplados desde los primeros cursos de primaria. Por tanto, podemos esperar que nuestros resultados fueran mejores.

Desde el análisis de la aleatoriedad desde el significado frecuencial de la probabilidad, se identifican investigaciones cuyo propósito es describir la comprensión de los alumnos sobre las características de las secuencias aleatorias. Desde esta perspectiva, Falk y Konold (1997) proporcionan una clasificación de las investigaciones enfocadas en el análisis de la comprensión de la aleatoriedad, considerando dos grupos: (a) Tareas de generación, las cuales se caracterizan por que el sujeto genera una serie de secuencias como resultados de un proceso aleatorio; y (b) Tareas de reconocimientos, en las que sujeto elige aquellas secuencias que considera aleatorias. De hecho las tareas que nosotros proponemos se pueden considerar como de generación, aunque usamos secuencias muy cortas (sólo cuatro resultados).

La investigación de Green (1983a) recoge una tarea constituida por dos sucesiones que simulan el resultado de lanzar una moneda equilibrada 150 veces. La tarea de reconocimiento demanda que el alumno identifique cuál de las dos secuencias es una secuencia aleatoria. El estudio muestra que la mayor parte de los niños escogen la secuencia no aleatoria. Por otro lado, el análisis de los argumentos dados por los alumnos para la no aleatoriedad conduce a una clasificación de los mismos en correctos e incorrectos. El autor identifica como correctos: patrón demasiado regular o rachas muy cortas, e incorrectos: patrón demasiado irregular, el porcentaje de caras no es 50% o rachas muy largas. En relación con nuestro trabajo, este estudio es relevante porque una secuencia de resultados puede verse como una muestra generada por un proceso.

Posteriormente, Green (1991) realiza una investigación sobre una muestra de 305 niños ingleses (7 a 11 años) donde compara las respuestas a las preguntas formuladas sobre secuencias aleatorias. Una de las preguntas formuladas se corresponde con una tarea de generación, donde se pide que escriba una sucesión de tamaño 50 que simulara el lanzamiento de una moneda. El autor lleva a cabo un seguimiento longitudinal del grupo de alumnos durante cuatro años a través del cual detecta que no existe una mejora durante esos años. Los resultados de la investigación reflejan que los niños tienden a subestimar la variabilidad en la secuencias. Además, escriben secuencias donde la proporción del número de caras y cruces es próxima $\frac{1}{2}$ con una dispersión más pequeña que la que se espera de una secuencia aleatoria.

Serrano (1996) diseñó un cuestionario de 10 ítems constituido por preguntas abiertas y de selección múltiple, que describen tareas de reconocimiento y generación de secuencias aleatorias, así como la interpretación frecuencial de la probabilidad. El cuestionario fue respondido por una muestra constituida por alumnos españoles, entre los que se distinguen 147 alumnos de 13 años, y 130 alumnos de 17 años. En relación con la primera de las tareas comentadas, asociada al reconocimiento de sucesiones aleatorias, el autor identificó tres concepciones: variabilidad local, regularidad global y grado de discrepancia entre la distribución esperada y la observada. Por otro lado, el autor identifica la ausencia de una comprensión adecuada de los siguientes aspectos: el uso de razones frecuenciales para explicar la ocurrencia de un suceso poco probable, la estimación de la frecuencia relativa de un suceso en una serie futura de experimentos cuando se conoce su probabilidad y la estimación frecuencial de una probabilidad no sirve para predecir la ocurrencia del resultado de un único experimento.

Más recientemente Gómez et al. (2014) elaboran un cuestionario constituido por cuatro ítems para evaluar algunas componentes del conocimiento matemático para la enseñanza de la probabilidad desde el enfoque frecuencial de futuros profesores de Educación Primaria. El primer ítem está dirigido a evaluar el conocimiento común del contenido (Ball, Lubienski y Mewborn, 2001), y es una adaptación del propuesto por Green (1983a). Además, será uno de los que usaremos en nuestro estudio. Dicho ítem demanda a los participantes pensar sobre la elaboración de cuatro muestras probables de un tamaño determinado. Los resultados se desprenden desde el análisis del contenido de los cuatro valores dados, centrado en la media de los cuatro valores y su variabilidad. En este caso, los autores identifican que sólo una tercera parte aproximadamente tiene una intuición simultánea de la convergencia al valor esperado y de la variabilidad muestral. Por tanto, identifican que los sujetos presentan pobres intuiciones iniciales sobre los experimentos aleatorios. Por otro lado, se identifica que los participantes presentan diferentes sesgos como el sesgo de equiprobabilidad, así como la heurística de la representatividad; o piensan que no es posible realizar la predicción.

Estos resultados coinciden con las investigaciones de Green (1983a, 1991), Serrano (1996) y Cañizares (1997), aunque son mejores. Por otro lado, Gómez et al. (2014) señalan que la adaptación del ítem proporciona información sobre la comprensión de la variabilidad. En este sentido, los resultados muestran que una parte de los sujetos produce muestras de variabilidad extrema o de patrón determinista.

2.7. CONCLUSIONES DEL ESTUDIO DE LOS ANTECEDENTES

En primer lugar, la revisión de las investigaciones centradas en identificar la imagen que los alumnos asocian al concepto de muestra revela que aparentemente los estudiantes visualizan las muestras como conjuntos excluyentes de la misma población y no como conjuntos con elementos comunes. Por otro lado, aunque identifican los tres niveles de distribución que aparecen en el muestreo: datos, muestra y población, no se observa que relacionen la distribución muestral, con la distribución de la población. Además, se identifica que son capaces de llevar a cabo los procedimientos desligados de una adecuada comprensión de los conceptos.

Desde el enfoque frecuencial de la probabilidad, el estudio de los antecedentes nos muestra que, respecto a la idea de muestra, se comprende bien la representatividad de la misma, pero no la variabilidad. En las distintas investigaciones se identifica que los

sujetos que componen la muestra, o bien no saben reconocer la variabilidad en las tareas de reconocimiento, o bien no la saben reproducir en las tareas de generación. No obstante, las tareas demandan la reproducción de secuencias de resultados aleatorios, con excepción del trabajo de Gómez et al. (2014), donde se demanda la reproducción de muestras. Esta diferencia coincidente con el propósito de nuestro trabajo conlleva la consideración del ítem en el instrumento de evaluación que diseñamos.

Al igual que en la mayoría de los trabajos descritos, el interés se centra en evaluar si los alumnos presentan una adecuada comprensión de la variabilidad muestral, así como del efecto del tamaño de la muestra en la misma. No obstante, la evaluación se sustenta en una serie de cuestiones en las que se demanda al sujeto la elaboración de muestras grandes y pequeñas en relación con el experimento descrito en el enunciado. A diferencia de las investigaciones citadas anteriormente, los enunciados que configuran el cuestionario que presentamos a continuación, hacen referencia a fenómenos aleatorios caracterizados por sucesos equiprobables o no. Esta particularidad permite comparar respuestas según este rasgo, siendo este punto objeto de análisis de este trabajo.

CAPÍTULO 3.

ESTUDIO DE EVALUACIÓN

3.1. INTRODUCCIÓN

Este capítulo presenta un doble objetivo. Por un lado, describir el instrumento diseñado para evaluar algunas de las ideas que se desprenden del concepto de muestreo, las cuales se citan en los objetivos 3 y 4 del primer capítulo. El análisis previo del cuestionario supone la caracterización de cada ítem, a través del razonamiento que guía en la determinación de su respuesta correcta asociada.

Por otro lado, se pretende analizar la comprensión que una muestra de estudiantes que participa en el estudio tiene sobre los conceptos asociados al muestreo. Para ello, se examinan las respuestas obtenidas por dicha muestra tras la aplicación del cuestionario. En particular, este análisis supone la reflexión sobre las respuestas de los estudiantes, identificando los posibles sesgos o concepciones erróneas acerca de los conceptos evaluados. Por tanto, se consideran algunos de los aspectos comentados durante el Capítulo 2 para el análisis de los datos correspondientes con las respuestas de la muestra de estudiantes.

El capítulo se inicia con la caracterización del contexto del estudio. A continuación, se resume la metodología utilizada, donde se recoge tanto la caracterización de la muestra como la descripción del cuestionario diseñado. Tras la aplicación del instrumento de evaluación, el capítulo se centra en analizar las respuestas elaboradas al cuestionario. Finalmente, se señalan los aspectos más relevantes que se obtienen desde el análisis de las respuestas para cada uno de los ítems. Además, se analizan los datos según las características o particularidades que presenta cada enunciado.

3.2. CONTEXTO EDUCATIVO

Aunque el Capítulo 1 presenta un análisis de la presencia de los contenidos asociados al muestreo a lo largo de la etapa educativa de la ESO, este trabajo está enfocado en analizar la comprensión de los alumnos de segundo y cuarto curso. Por tanto, el contexto educativo está centrado en la etapa obligatoria de secundaria, y en particular, en esos dos cursos.

Los participantes que configuran la muestra pertenecen a dos institutos (centros públicos) que se localizan en la ciudad de Huesca. Aunque los estudiantes provienen de dos centros distintos, las diferencias debidas a este hecho no son relevantes puesto que diversos factores como, la situación socio-económica de las familias es muy similar en ambos institutos.

El acceso a los estudiantes, que constituyen la muestra, fue facilitado por los directores de los centros y por la Consejería de Educación, a los que se solicitó permiso para realizar el estudio. Igualmente, se pidió la colaboración a los profesores de los grupos, los cuales la prestaron voluntariamente y se interesaron por obtener los resultados finales del estudio. De tal forma que agradecemos nuevamente a todos su cooperación, sin la cual no hubiese sido posible realizar esta investigación.

3.3. METODOLOGÍA

En cuanto a la metodología seguida, el trabajo realizado se trata de un estudio exploratorio de evaluación, en el que el instrumento o herramienta de evaluación diseñado toma el primer ítem de otro trabajo de investigación, mientras que el resto son de elaboración propia. Además, señalar que el primer ítem se aplicó a futuros profesores de educación primaria, de mayor edad y formación que nuestros estudiantes.

En síntesis, el estudio se corresponde con una investigación preferentemente cuantitativa, puesto que se basa en el análisis estadístico de los datos obtenidos a partir de las respuestas dadas por los estudiantes; pero se complementa con algunos aspectos cualitativos, porque analizamos algunos tipos de respuestas que indican sesgos de razonamiento. Según las clasificaciones de Bisquerra (1989), se trata de una investigación aplicada y descriptiva. La muestra de alumnos es intencional, aunque elegida de forma que represente adecuadamente a otros alumnos del mismo contexto.

El análisis de datos es descriptivo y se reduce a la elaboración y comparación de gráficos estadísticos. En particular, se exponen tanto las tablas de frecuencias como los diagramas de cajas asociados a los valores medios y el rango obtenidos desde las muestras dibujadas por los sujetos. Para realizarlo se codificaron previamente los datos y se utilizó el paquete estadístico SPSS.

3.3.1. DESCRIPCIÓN DE LA MUESTRA

La muestra está constituida por un total de 302 alumnos, de los cuales 157 son de 2ºESO (12-13 años) y 145 de 4ºESO (15-16 años). El número de grupos distintos de alumnos que participaron fue 9 grupos de 2ºESO, y 8 grupos de 4ºESO, de los cuales 6 grupos corresponden a la opción B.

En concreto, el cuestionario es respondido por 108 alumnos que cursan la opción B (74,5% de los alumnos de 4ºESO). Siguiendo las directrices curriculares, se considera que la opción A está enfocada para aquellos alumnos que quieran realizar una formación profesional. Mientras que la opción B está dirigida a los estudiantes que pretenden continuar con los estudios de Bachillerato. En relación con los contenidos que el currículo refleja para la opción B, observamos que la diferencia con respecto a la primera opción anterior supone la consideración de algunos contenidos nuevos.

Los alumnos de 2ºESO no habían recibido instrucción previa sobre contenidos de probabilidad y estadística, aparte de los conocimientos que pudieran haber obtenido en la Educación Primaria, donde en el último curso hay algunas ideas intuitivas sobre sucesos aleatorios y probabilidad simple, pero no se estudia el muestreo.

Algunos de los alumnos de los grupos de 4ºESO han recibido instrucción sobre probabilidad desde el significado clásico de la misma o nociones de estadística descriptiva en los cursos anteriores (ver contenidos en el análisis curricular), independiente de la modalidad que están cursando actualmente. Por su parte, la ausencia de enseñanza de dichos contenidos responde a factores como el tiempo, que imposibilita a los profesores poder llevar al aula todos los contenidos que cita el currículo. En el caso de la enseñanza de la probabilidad y la combinatoria, dichos contenidos están presentados al final del temario, lo cual conduce a la observación citada anteriormente, es decir, que dicho contenido no sea enseñado.

3.3.2. MÉTODO DE RECOGIDA DE DATOS

Los datos fueron recogidos en todos los grupos participantes en horas en las que los alumnos tenían clases de matemáticas, como una actividad de la citada clase. Los estudiantes completaron el cuestionario por escrito en un intervalo de tiempo de 15-25 minutos. Durante la aplicación del cuestionario, el docente responsable del grupo aula

estaba presente, aunque situado en un segundo plano. Es decir, el investigador es la persona que describe las indicaciones y finalidad del cuestionario.

Se les explicó a los estudiantes el fin de la evaluación y se resolvieron posibles dudas. Adicionalmente, se observa que la mayoría de los alumnos respondieron con interés, completando su mayoría todas las preguntas.

Tras la obtención de los datos, se ha procedido a la codificación de los mismos asignando un código a cada cuestionario, constituido por los siguientes símbolos: número del curso (2 ó 4), en el caso de que el alumno sea de 4ºESO, se indica la modalidad que está cursando, A o B. Luego, se considera la inicial del instituto al que pertenece el alumno. Finalmente, se indica el número que el sujeto ocupa en la muestra obtenida.

3.4. ANÁLISIS A PRIORI DEL CUESTIONARIO Y CONOCIMIENTOS EVALUADOS

Para el diseño del cuestionario se han considerado algunos ítems utilizados en investigaciones previas o, adaptados de la mismas. De hecho, la primera tarea es una adaptación de una tarea utilizada en el trabajo de investigación de Green (1983a), que fue diseñada para evaluar la comprensión de la estimación frecuencial. Además, dicho ítem fue adaptado por Gómez et al. (2014), para poder obtener información sobre la comprensión de la variabilidad en el muestreo. Entonces, en nuestro trabajo se ha utilizado directamente el ítem propuesto por Gómez et al. (2014). Por otro lado, el diseño del resto de las tareas que constituyen el cuestionario, son variantes del primero; en todas ellas se demanda al alumno escribir cuatro muestras asociadas a un experimento concreto descrito en el enunciado.

De manera genérica, el cuestionario está constituido por cuatro ítems que tienen como finalidad evaluar tanto la comprensión de la variabilidad del muestreo como el efecto del tamaño de la muestra sobre dicha variabilidad. Además, esta herramienta permite observar si el alumno muestra la relación existente entre el valor de una proporción en la población y la frecuencia relativa esperada en una muestra. A continuación, se describe detalladamente cada una de las tareas del mismo, razonando sobre su solución correcta, la cual nos permite caracterizar si el estudiante presenta una comprensión adecuada de los conceptos evaluados.

3.4.1. ANÁLISIS DEL ÍTEM 1

Item 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:

Daniel	Martín	Diana	María
Punta arriba:	Punta arriba:	Punta arriba:	Punta arriba:
Punta abajo:	Punta abajo:	Punta abajo:	Punta abajo:

Este ítem, se ha tomado de la investigación de Gómez et al. (2014), como se ha comentado anteriormente. El fenómeno aleatorio que formula el enunciado se corresponde con el lanzamiento de chinchetas, entonces la distribución del número de chinchetas que caen con la punta hacia arriba se corresponde con un modelo binomial, $B(n, p)$, donde n es el tamaño de la muestra. Entonces, definimos la variable X como el número de chinchetas que caen hacia arriba, donde n toma el valor 100 y p es la probabilidad de que una chincheta caiga con la punta hacia arriba, siendo esta probabilidad desconocida. No obstante, se puede estimar a partir de los datos dados en el enunciado, los cuales permiten también obtener una estimación tanto del valor esperado, np , como de la desviación típica, $\sqrt{np(1 - p)}$. Para este caso, el valor esperado es estimado por 68, y su desviación típica es 4,7. Por tanto, se considera 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. Se considerarán cercanas aquellas respuestas cuyos valores, para el número de chinchetas que caen con la punta hacia arriba, se localicen dentro del intervalo (63,73), que se calcula al determinar la media más/menos la desviación típica. Este es el intervalo de valores que contiene el 68% de las observaciones en la distribución normal. Finalmente, otro aspecto que se debe reconocer es la ausencia de equiprobabilidad por la asimetría física del dispositivo.

Con respecto al análisis de la comprensión de la variabilidad se estudia previamente el rango obtenido a partir de los valores, es decir, el valor máximo menos el mínimo proporcionados para la frecuencia absoluta asociada al número de chinchetas que caen con la punta hacia arriba. Como se ha indicado, la desviación típica en este experimento es aproximadamente igual a 5; por tanto se considera el intervalo que se determina al considerar dos veces la desviación típica, puesto que contiene el 95% de los valores. Por tanto, los valores de las chinchetas que caen con la punta hacia arriba

deberían presentar una variabilidad entre 58 y 78, que se corresponde con un rango asociado a los cuatro valores proporcionados de una amplitud de 20. Siguiendo a Gómez et al. (2014), se considera adecuada la variabilidad cuando el rango se encuentra en un intervalo la mitad del cual tiene entre una y dos desviaciones típicas de longitud, aproximadamente entre 10 y 20 (una y dos desviaciones típicas sumadas y restadas a la media). Si está entre dos y tres veces la desviación típica (entre 20 y 30) se considera alta, pero aceptable y si es mayor excesiva. Por otro lado, si es menor que 10 se considera demasiada concentración y quiere decir que no se comprende la variabilidad muestral.

3.4.2. ANÁLISIS DEL ÍTEM 2

Item 2. Un profesor vacía sobre la mesa un paquete de 100 monedas obteniendo los siguientes resultados: 53 caen con la cara hacia arriba y 47 caen con la cruz hacia arriba. Supongamos que el profesor pide a 4 niños repetir el experimento. Cada niño lanza las 100 monedas y obtendrá algunas con la cara hacia arriba y otras con cruz hacia arriba.

Escribe en la siguiente tabla un resultado que te parezca probable para cada niño:

Elena	Clara	Matías	Rosa
Cara:	Cara:	Cara:	Cara:
Cruz:	Cruz:	Cruz:	Cruz:

El fenómeno aleatorio descrito en este ítem se corresponde con la repetición del lanzamiento de una moneda, donde el número de ensayos es 100. Entonces, la variable aleatoria definida como el número de monedas que caen con la cara hacia arriba sigue una distribución binomial, donde n toma el valor 100. Desde el significado clásico de la probabilidad se tiene que los dos resultados del experimento (obtener cara o cruz) son equiprobables, es decir, la probabilidad de que salga cara es $\frac{1}{2}$. Tras la caracterización de la distribución que sigue la variable aleatoria, el cálculo de la desviación típica da un valor, $\sigma = 5$. Por tanto, se consideran como correctas aquellas respuestas cuyo promedio asociado a los cuatro valores para el número de caras se acerque a 50 y oscile dentro del intervalo (45,55).

Para el análisis de la comprensión de la variabilidad se estudia el rango asociado a los valores escritos, que se obtiene de calcular el valor máximo menos el mínimo proporcionados. Razonando de manera análoga al ítem anterior, se considera adecuada la variabilidad cuando el rango se encuentra entre una y dos desviaciones típicas, aproximadamente entre 10 y 20. Si está entre dos y tres veces la desviación típica (entre 20 y 30) se considera alta, pero. Finalmente, si es menor que 10 se considera demasiada

concentración mientras que si es mayor que 30 es excesiva. En ambos casos se identifica que el sujeto no presenta una comprensión adecuada de la variabilidad muestral.

3.4.3. ANÁLISIS DEL ÍTEM 3

Item 3. Un profesor pide a 4 niños lanzar 10 monedas sobre la mesa y contar el número de caras y cruces obtenidos. Escribe en la siguiente tabla un resultado que te parezca probable para cada niño:

Silvia	Javier	Miguel	Carmen
Cara:	Cara:	Cara:	Cara:
Cruz:	Cruz:	Cruz:	Cruz:

El fenómeno aleatorio que describe este enunciado se corresponde con el presentado en el ítem anterior. La diferencia radica en que se realizan 10 ensayos del experimento aleatorio. Entonces, la variable aleatoria sigue una distribución binomial, $B(10, 1/2)$. Por tanto, se considera que una respuesta es correcta, si la media de los cuatro valores toma un valor en torno al valor medio esperado y oscila en el intervalo (3,4, 6,6). Por otro lado, se considera adecuada la variabilidad cuando el rango se encuentra entre una y dos desviaciones típicas, aproximadamente entre (3,2, 6,4). Si está entre dos y tres veces la desviación típica (6,4, 9,6) se considera alta, pero aceptable y si es mayor excesiva. Si es menor que 2 se considera demasiada concentración, y si es mayor que 9, es demasiado alta. Estos dos últimos casos implican que no se comprende la variabilidad intrínseca al proceso de muestreo.

Como se ha citado, el fenómeno aleatorio es análogo al ítem anterior, puesto que se corresponde con lanzar una moneda. La diferencia se debe al número de ensayos para los que se repite el experimento. Este hecho permite analizar el efecto del tamaño de la muestra en la variabilidad de las muestras elaboradas para cada una de las tareas. Por tanto, el análisis comparativo de los dos ítems permite estudiar si existe una adecuada comprensión o, intuición de la Ley de los Grandes Números.

3.4.4. ANÁLISIS DEL ÍTEM 4

Item 4. Un jugador de baloncesto suele encestar 70 de cada 100 tiros a una canasta desde la posición de tiros libres. Escribe en la siguiente tabla un resultado que sea probable para cuatro partidos en los que lanza 10 tiros desde el punto de lanzamientos personales.

Partido 1 (10 tiros)	Partido 2 (10 tiros)	Partido 3 (10 tiros)	Partido 4 (10 tiros)
Número de encestes:	Número de encestes	Número de encestes	Número de encestes
Número de fallos:	Número de fallos:	Número de fallos:	Número de fallos:

La tarea describe la frecuencia de encestes de un jugador de baloncesto. La elección de esta situación se debe a la familiaridad del contexto, lo cual puede ayudar al alumno en la resolución del ítem. Siguiendo un razonamiento similar a los ítems anteriores, la variable aleatoria queda definida por el número de encestes, la cual sigue una distribución binomial $B(10, p)$, donde p es desconocida. Si consideramos el dato del enunciado, se estima que el valor esperado es 7, y su desviación estándar es 1,45. Por tanto, el valor esperado estimado asociado a los cuatro valores escritos para el número de encestes puede oscilar en el intervalo (5,55, 8,45), redondeando, entre 6 y 9.

Como en el ítem anterior, se considera adecuada la variabilidad cuando el rango se encuentra entre una y dos desviaciones típicas, aproximadamente entre 3 y 6. Si está entre dos y tres veces la desviación típica (entre 6 y 9 redondeando) se considera alta, pero aceptable y si es mayor excesiva. Si es menor que 3 se considera demasiada concentración y, quiere decir que no se comprende la variabilidad muestral.

En resumen, el instrumento de evaluación está constituido por cuatro ítems, la Tabla 3.4.1 sintetiza las características de los ítems, donde dos de ellos demandan la producción de muestras grandes (1 y 2) y otros dos, pequeñas (3 y 4). Además, en dos de los ítems (1 y 4) la asignación de probabilidad se hace de forma frecuencial, mientras que los otros dos (2 y 3) se apoyan en el significado clásico de la probabilidad a través de la regla de Laplace. Finalmente, los fenómenos aleatorios que se describen en el ítem 1 y 2, a diferencia de los otros dos, no presentan el principio de equiprobabilidad.

Tabla 3.4.1. Características de los ítems del cuestionario

Item	Tamaño de la muestra	Probabilidad de éxito	Asignación probabilidad	Contexto
1	100	0,68	Frecuencial	Chinchetas
2	100	0,5	Clásica o frecuencial	Monedas
3	10	0,5	Clásica	Monedas
4	10	0,7	Frecuencial	Encestes

3.5. RESULTADOS

Los siguientes puntos muestran los resultados asociados a cada una de los ítems que constituyen el cuestionario. En primer lugar, el estudio se centra en la muestra total

de estudiantes que responden el cuestionario. Tras este análisis previo, el interés radica en identificar las posibles diferencias en función del curso con el propósito de estudiar si existe un progreso con la edad y/o enseñanza recibida.

Por un lado, el análisis recae inicialmente sobre el estudio de la distribución asociada al valor medio de las cuatro muestras dadas por el alumno. Este estudio permite observar si las muestras proporcionadas tienen una frecuencia relativa esperada adecuada en relación con el valor que toma la proporción en la población, para lo cual se considera el histograma de frecuencias asociado a dichas medias. Este gráfico permite observar si la muestra participante proporciona frecuencias cuyo valor medio se localiza en el intervalo que se ha denominado como aceptable. Por otro lado, el análisis de la variabilidad esperada por los alumnos durante el proceso de muestreo recae sobre el estudio del rango asociado a los cuatro valores proporcionados, el cual se ha definido en el apartado anterior. Análogamente, el estudio queda fundamentado en el histograma de frecuencias asociado al rango.

Finalmente, este estudio previo de la muestra se complementa, como se ha señalado, con un estudio comparativo de los cursos participantes. Para lo cual, se considera, además, el diagrama de caja asociado al valor medio y al rango, respectivamente. Esta representación permite observar la mediana y los cuartiles y la simetría de la distribución. Adicionalmente, este gráfico señala la existencia de valores atípicos, hecho que favorece la identificación de respuestas asociadas a un sesgo particular.

3.5.1. RESULTADOS EN EL ÍTEM 1

El análisis previo de las respuestas refleja que se han descartado 17 respuestas (10 correspondientes a alumnos de segundo curso y 7 de cuarto) por razones diversas, entre las que destacan dos principalmente: el alumno no responde o proporciona una respuesta inadecuada por diversos motivos, por ejemplo, escriben valores que no pertenecen al rango de respuestas posibles. Estas causas se identifican durante el análisis de las respuestas de los distintos ítems, por lo que no se cita durante el análisis de los siguientes ítems, ya que dicha información queda recogida en la tabla elaborada para el estudio del valor medio.

En este ítem, la única información que se proporciona al estudiante (frecuencial) es el número de chinchetas que el profesor obtuvo con la punta hacia arriba al realizar el experimento. Entonces, los alumnos pueden razonar, desde su experiencia, que es más probable que la chincheta caiga hacia arriba, debido al peso de la cabeza.

En relación con el análisis de la estimación sobre la frecuencia esperada para el número de chinchetas que caen con la punta hacia arriba, se considera la distribución del valor medio asociada a las cuatro respuestas proporcionadas por el estudiante, cuya representación se expone en la Figura 1a. Desde el análisis de dicha distribución, se obtiene un valor medio de 57,88 (resultado muy similar al obtenido en el trabajo de Gómez et al. (2014), donde el valor medio de dicha distribución es 57,7). Sin embargo, este valor es algo inferior al valor medio esperado, que es 68 (línea vertical señalada la Figura 1a). En definitiva, los estudiantes presentan una estimación del número medio de chinchetas que caen hacia arriba sesgada hacia el valor 50, debido a que algunos no tienen en cuenta la información frecuencial que proporciona el enunciado.

Además, se observan dos picos en la distribución. Por un lado, se identifica que el 24,50% de la muestra proporciona valores medios que se localizan en el intervalo (45, 55). Estos estudiantes presentan el sesgo de equiprobabilidad (Lecoutre, 1992), consistente en considerar como equiprobables resultados de un experimento aleatorio que claramente no lo son, como el caso de la chincheta. De los cuales solo en un 3,97% elaboran respuestas en las que existe ausencia de variabilidad total, es decir, escriben una cuaterna (50, 50, 50, 50). Estas respuestas también aparecen en el trabajo de Gómez et al. (2014). Por otro lado, si analizamos la frecuencia de los valores medios próximos al valor esperado (63,73), este grupo supone el 32,45%; serían los estudiantes que proporcionan una buena estimación del valor esperado en el experimento. Otro grupo (13,91%) muestra una tendencia a dar valores medios más bajos del 50%, es decir, consideran que la proporción de chinchetas que caen con la punta hacia arriba es menor. Esta tendencia se traduce en la intención de compensar el resultado dado en el enunciado, mostrando la heurística de la representatividad (Tversky y Kahneman, 1982).

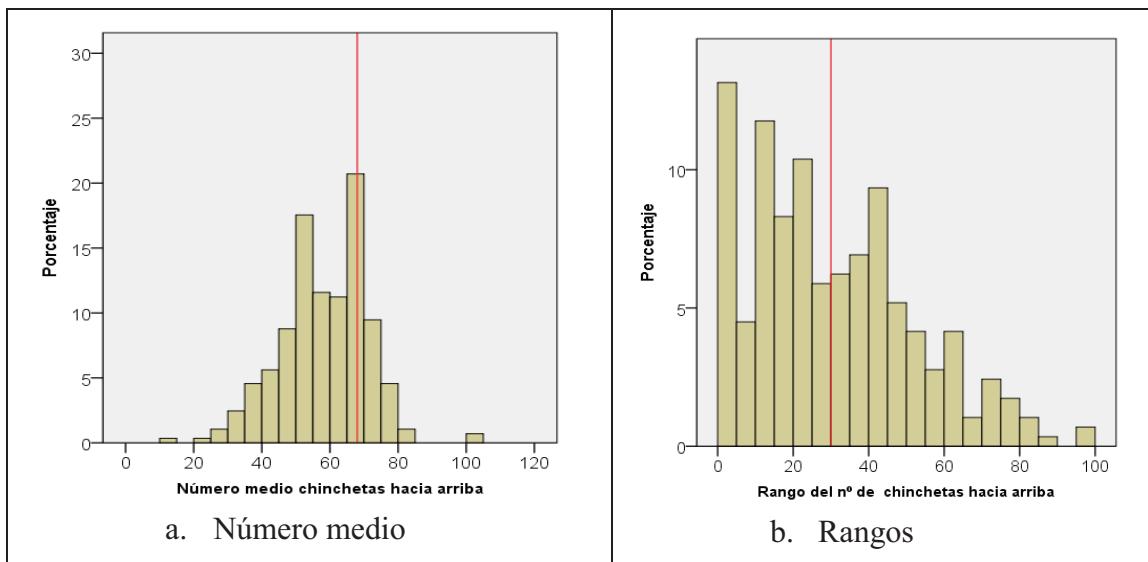


Figura 1. Distribución del número medio y rango de 4 valores del nº de chinchetas con la punta hacia arriba

Por otro lado, la Figura 1b muestra la distribución de rangos asociada al conjunto de las cuatro respuestas de cada estudiante, la cual permite analizar la comprensión de la variabilidad del muestreo en estos estudiantes. Como se ha indicado anteriormente, y siguiendo a Gómez et al. (2014) se considera adecuada la variabilidad cuando el rango se encuentra entre una y dos desviaciones típicas, aproximadamente entre 10 y 20, alta, pero aceptable si está 20 y 30 y si es mayor excesiva. Si es menor que 10 se considera demasiada concentración. El análisis de las frecuencias con la que se presentan los rangos en dichos intervalos conduce a observar que la mayoría produce una variabilidad excesiva. En particular, se identifica que el 40,4% proporciona respuestas con una dispersión muy alta (mayor que 30) para una muestra del tamaño dado. Además, el 20,8% proporciona datos con apenas variabilidad (rango menor que 10), mientras el resto dan una variabilidad adecuada o alta, pero razonable. Por tanto, son muchos los estudiantes que no comprenden el efecto del tamaño de muestra sobre la variabilidad del muestreo, en línea con las investigaciones de Serrano (1996) y Shaughnessy et al. (2004).

Comparación por curso

A continuación el interés radica en analizar las respuestas según el curso participante, a través de las Figuras 2 y 3, respectivamente. Dicha información se complementa con las Tablas 1 y 2, las cuales presentan la frecuencia y porcentaje de

alumnos según determinados intervalos, tanto para el estudio del valor medio como para el análisis del rango.

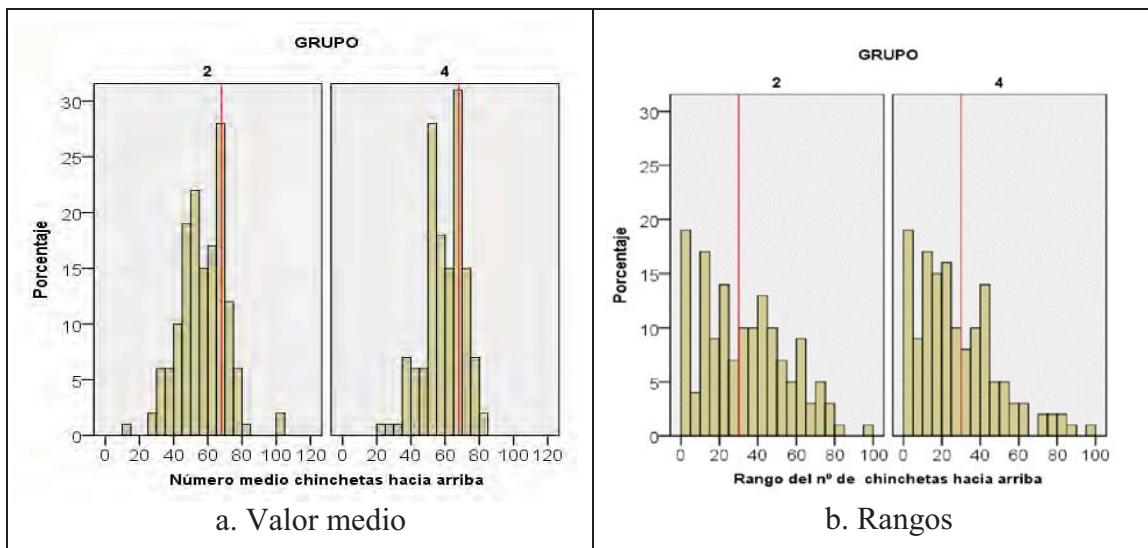


Figura 2. Distribuciones del valor medio y rango del nº de chinches con la punta hacia arriba según el curso

En primer lugar, el valor medio asociado al grupo de 2ºESO (57 aprox.) es menor que el de 4ºESO (61 aprox.). Asimismo, el análisis de la distribución del valor medio presentada en la Figura 2a, evidencia que en ambos grupos los alumnos no identifican la simetría del dispositivo. Este hecho puede corresponderse a la presencia tanto del sesgo de equiprobabilidad como de la heurística de la representatividad, citados anteriormente. Además, esta información se completa con la información resumida en la Tabla 1, donde se observa la frecuencia con que aparecen tanto el sesgo de equiprobabilidad y como 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), siendo la presencia del primero más relevante.

Por otro lado, en este estudio se identifica que la proporción de sesgos, así como de respuestas inadecuadas o en blanco es mayor en el 2º curso, aunque con pequeñas diferencias.

Tabla 1. Frecuencia y porcentaje de alumnos según intervalo en que se sitúa la media de los valores dados

	2ºESO		4ºESO	
	Frecuencia	%	Frecuencia	%
Hasta 45- Representatividad	26	16,56	16	11,03
45-55. Equiprobabilidad	40	25,48	34	23,45
63-73 Correcto	46	29,30	52	35,86
Otros valores aceptables	35	22,30	36	24,83
Otros casos ¹	10	6,40	7	4,80

En segundo lugar, se realiza un análisis similar para el estudio de la variabilidad. Por un lado, la Figura 2b representa la distribución del rango según el curso. Desde su análisis, se identifica que ambos grupos conceden una variabilidad extrema a las muestras proporcionadas. De hecho, se identifica que, aproximadamente, el 50% de los alumnos de 2ºESO que participan en el estudio muestran una comprensión pobre acerca de la variabilidad asociada, siendo este porcentaje menor que en el otro grupo estudiado (ver Tabla 2). No obstante, ambos grupos elaboran respuestas con un variabilidad extrema con alta frecuencia. Además, se identifica que el siguiente rango con mayor presencia se corresponde con valores menores que 10, lo cual implica una concentración elevada de los valores escritos para el número de chinchetas que caen con la punta hacia arriba.

Tabla 2. Frecuencia y porcentaje de alumnos según intervalo en que se sitúa el rango de los valores dados

	2ºESO		4ºESO	
	Frecuencia	%	Frecuencia	%
Hasta 10. Alta concentración	30	19,11	27	18,62
10-20. Adecuada	20	12,74	29	20
20-30. Alta (pero aceptable)	21	13,36	26	17,93
Mayor que 30. Excesiva	77	49,04	56	38,62

Finalmente, el análisis del diagrama de cajas asociado al valor medio en cada grupo (Figura 3a), señala, por un lado, que la mediana es algo más baja en los alumnos de segundo. Pensamos que esto ocurre porque con mayor frecuencia estos alumnos dan valores cercanos al 50%; es decir, los alumnos entienden que los dos resultados son equiprobables, en vez de usar la información frecuencial. Por otro lado, la dispersión

¹ Bajo esta categoría se engloban tanto las respuestas de los alumnos que son incorrectas por diferentes razones, como aquellos alumnos que no responden al ítem.

asociada a este estadístico es similar en los dos grupos. Los valores atípicos identificados se corresponden con respuestas cuyos valores para el número de chinchetas se sitúa en los valores extremos del rango de respuestas posibles, como se observa en las siguientes respuestas de los estudiantes: 2P_A61 y 2S_A135: (100, 100, 100, 100), 2P_A24: (22, 0, 8, 16) y 4BP_A216: (20, 40, 20, 15). 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.

La parte derecha del gráfico (Figura 3b), que corresponde a los rangos de las 4 puntuaciones, muestra menor mediana y menor tercer cuartil en el grupo de alumnos de 4º curso, confirmando que éstos parecen conceder en mayor proporción una variabilidad menor a las muestras cuyos tamaño es grande. Análogamente, se identifican algunos valores atípicos, entre los que destaca la siguiente respuesta, 4BS_A263: (99, 1, 50, 60). Ésta revela que el estudiante tiene una concepción errónea de la variabilidad muestral asociada al muestreo, puesto que no tiene en cuenta el efecto del tamaño de la muestra, lo que conduce a la aceptación como una respuesta probable todas aquellas respuestas que se localizan en el rango de valores posibles.

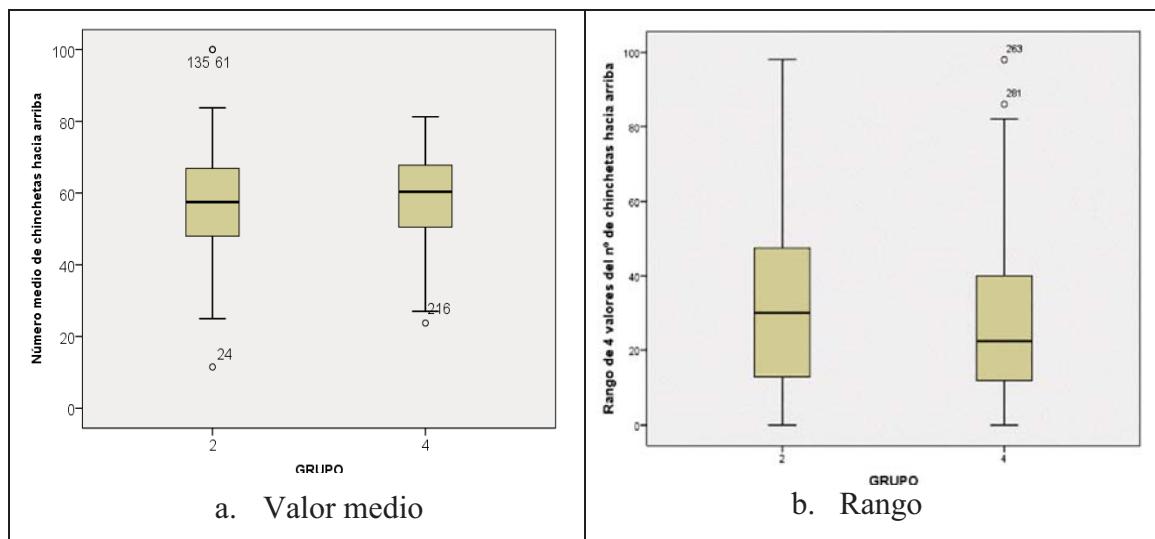


Figura 3. Diagramas de caja del número medio y rango de 4 valores del nº de chinchetas con la punta hacia arriba

3.5.2. RESULTADOS EN EL ÍTEM 2

De manera análoga al análisis realizado en el primer ítem, el estudio del segundo enunciado se centra en el análisis de la distribución del valor medio asociado a los

cuatro ensayos dados por el estudiante (ver Figura 4a). En primer lugar, se obtiene que el valor medio del estadístico, la media muestral, toma el valor de 51,21. El análisis previo del ítem refleja que el valor medio esperado es 50. Por tanto, se observa que el valor medio teórico para este ítem es próximo al valor medio que toma el estadístico en el muestreo. En consecuencia, los participantes muestran una adecuada estimación de la proporción del número de caras esperada en un lanzamiento de 100 monedas. Asimismo, los estudiantes muestran una comprensión adecuada de la estimación frecuencial. Quizás este resultado se corresponde con el hecho de que los estudiantes entienden y aplican el principio de equiprobabilidad, dando lugar a un porcentaje mayor de respuestas correctas en relación con lo obtenido para el primer ítem.

Desde el análisis de las frecuencias de los valores medios que se presentan en ciertos intervalos (ver Figura 4a y Tabla 3), se identifica, en primer lugar, que el 48,34% de los alumnos da valores próximos al valor esperado (45,55). Por tanto, proporcionan muestras tales que el valor medio del número de caras esperadas cae en el intervalo que se ha identificado como correcto.

Por otro lado, el análisis de las frecuencias que toman los distintos intervalos simétricos resulta menos llamativo. Entonces, interpretamos que los estudiantes identifican la equiprobabilidad del fenómeno aleatorio.

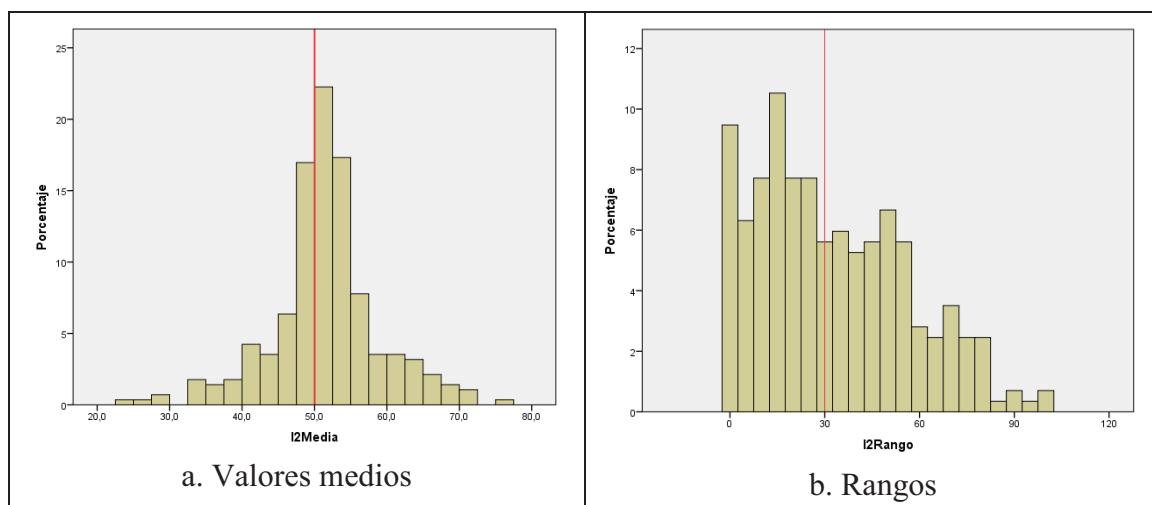


Figura 4. Distribuciones del valor medio y rango del nº de caras en 100 lanzamientos de monedas

Por otro lado, la Figura 4b representa la distribución del rango asociado a la cuaterna de muestras proporcionadas. En primer lugar, desde su análisis se identifica que los estudiantes conceden una variabilidad excesiva a las respuestas aportadas. De

hecho, la mitad de la muestra aproximadamente (46,69%) escribe respuestas cuyo rango excede de 30, que corresponde con una dispersión excesiva. Por tanto, razonando de manera análoga que en el ítem anterior, los estudiantes no comprenden el efecto del tamaño de la muestra, y producen valores con variabilidad excesiva, tal y como se identifica en la investigación de Serrano (1996) y Shaughnessy et al. (2004). Por otro lado, el 20% de la muestra participante produce valores cuyo rango es menor que 10, es decir, producen muestras con una concentración alta.

Comparación por curso

En primer lugar, se considera como objeto de estudio la distribución del valor medio según el grupo participante. El análisis comparativo de los dos grupos (ver Figura 5a y 6a) revela que la dispersión de los valores medios es más acusada en la muestra de los alumnos de 2ºESO, resultado que se observa también en el ítem anterior. Esta información se completa con los datos recogidos en la Tabla 2, donde se identifica que el porcentaje de alumnos que dan un valor medio correcto supone el 69% de los estudiantes de 4ºESO, mientras que dicho porcentaje es del 49,04% para el otro grupo. Por tanto, las respuestas de alumnos de 4ºESO reflejan, en media, una concentración en torno al valor teórico esperado. Sin embargo, el estudio de la variabilidad a través del cálculo del rango, indica que los estudiantes de 4ºESO conceden respuestas con una dispersión muy alta.

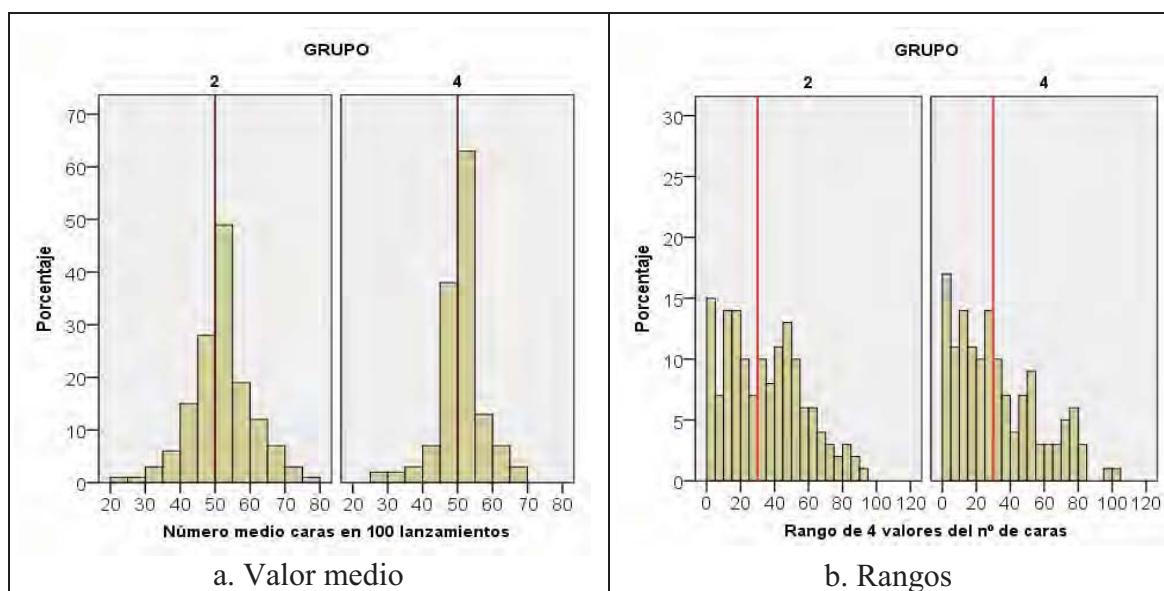


Figura 5. Distribuciones del valor medio y rango del nº de caras en 100 lanzamientos de monedas según curso

Tabla 3. Frecuencia y porcentaje de alumnos según intervalo en que se sitúa la media de los valores dados

	2ºESO		4ºESO	
	Frecuencia	%	Frecuencia	%
Correcto (45-55)	77	49,04	100	69
Menor de 45	27	17,20	15	10,34
Mayor de 55	42	26,75	23	15,86
Otros casos	11	7,01	7	4,80

Por su parte, la mitad de los estudiantes de 2ºESO proporcionan muestras cuyo valor medio es correcto. No obstante, de manera análoga al otro grupo, la variabilidad de las muestras dibujadas es muy alta. De hecho, la mitad de la muestra concede un rango mayor que 30. Este hecho conlleva a determinar que los estudiantes no presentan una intuición simultánea de las dos ideas antagónicas cuya coordinación es necesaria para una comprensión adecuada del concepto de muestreo. Como se ha señalado en el Capítulo 2, esas ideas fundamentales son la representatividad muestral y la variabilidad muestral (Batanero et al., 1994). Esta observación indica que la muestra participante muestra, en general, una comprensión parcialmente correcta de la Ley de los Grandes Números, resultado que se identifica también en el trabajo de Gómez et al. (2014). La comprensión es algo mejor en los alumnos de 4º, según muestran los datos.

Tabla 4. Frecuencia y porcentaje de alumnos según intervalo en que se sitúa el rango de los valores dados

	2ºESO		4ºESO	
	Frecuencia	%	Frecuencia	%
Hasta 10. Alta concentración	26	16,56	31	21,38
10-20. Adecuada	25	15,92	21	13,38
20-30. Alta (pero aceptable)	17	10,83	24	15,29
Mayor que 30. Excesiva	79	50,32	62	39,40

Finalmente, el estudio de los diagramas de caja (ver Figura 6) muestran menos dispersión para los valores medios en el grupo de cuarto, puesto que dan respuestas más homogéneas (ver Figura 6a). En relación a los valores atípicos identificados, éstos se corresponden con casos en los que las respuestas presentan una variabilidad muy baja, de modo que las respuestas asociadas se corresponde con la verosimilitud que les concede el alumno en función o no a su información previa. Las respuestas señaladas

son: 2S_A142: (70, 78, 75, 82), así como la respuesta del alumno 4BP_A174: (30, 25, 32, 18). Ninguno de estos conjuntos de datos representa, además, el valor esperado.

En el caso del estudio del rango (Figura 6b), se identifica como atípica la siguiente respuesta, 4BS_A263, (1, 99, 49, 51). Desde el análisis de los valores, se identifica que el estudiante no presenta una adecuada comprensión de la variabilidad muestral, donde se observa que el criterio que puede guiar en la determinación de una estimación probable para la proporción del número de caras es que pertenezca al intervalo [0, 100].

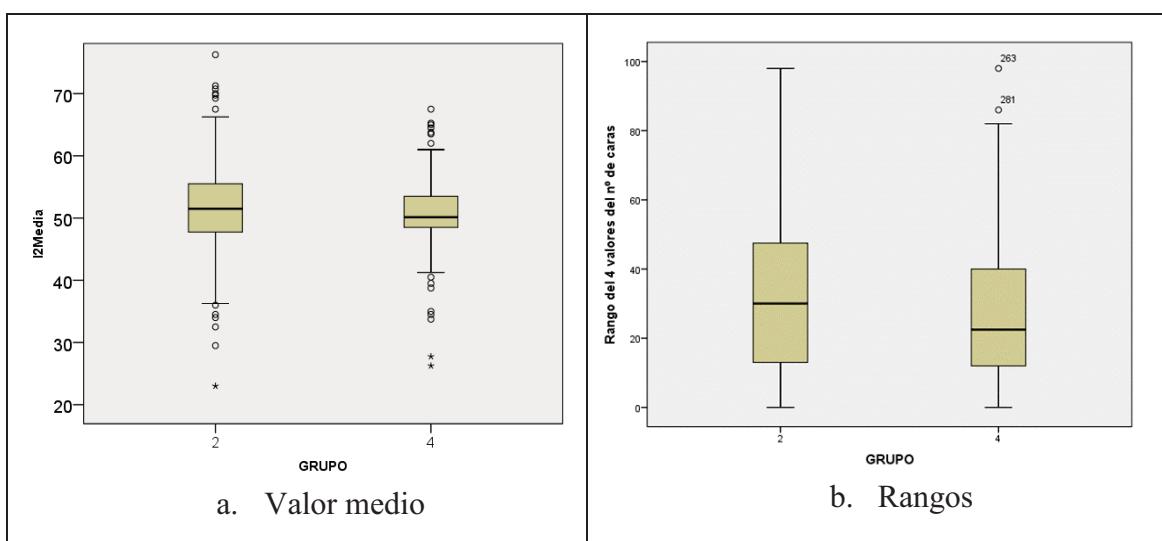


Figura 6. Diagramas de caja del valor medio y rango del nº de caras en 100 lanzamientos de monedas según curso

3.5.3. RESULTADOS ÍTEM 3

Como en los ítems anteriores, en primer lugar, se considera la distribución de las medias de los cuatro valores escritos por el alumno para el número de caras. Desde un estudio estadístico se calcula que el valor medio asociado a la muestra participante es 5,12. Además, este valor es muy próximo al valor esperado, 5, lo que indica una estimación buena del valor esperado en el global de la muestra. Este resultado se refuerza puesto que la desviación típica del estadístico es 1,58. Por tanto, los estudiantes proporcionan muestras que en media están concentradas en torno al valor medio teórico.

Desde el análisis de las frecuencias con que los valores medios se localizan en ciertos intervalos se infieren diferentes tipos de razonamientos. El análisis previo de la tarea refleja que las respuestas correctas son aquellas que consideran que el número de

caras toma un valor medio que varía dentro del intervalo (3,4, 6,6). En nuestro caso, se identifica que el 83,77% de los estudiantes proporcionan muestras donde la proporción del número de caras esperada se localiza en dicho intervalo. Por tanto, se observa que los estudiantes reflejan una buena concepción de la frecuencia esperada para este ítem. Dentro de este grupo de estudiantes, se identifica un subgrupo, 65,61%, que proporcionan valores medios entre 4,5 y 5,5. Este grupo proporciona unos resultados muy próximos al valor teórico, bien porque dan valores muy próximos al valor esperado o porque se compensan las distintas frecuencias.

De manera análoga al ítem anterior, se observa que los estudiantes identifican la equiprobabilidad del experimento, debido a la simetría que adopta la distribución del valor medio (Figura 7a).

En cuanto al estudio de la variabilidad, el análisis se sostiene en el rango de las muestras elaboradas por cada alumno. Por un lado, la Figura 7b recoge la distribución asociada a dicho estadístico en el muestreo, donde se identifica que la mayoría de los estudiantes proporcionan muestras cuyo rango varía dentro del intervalo (3, 5). Por tanto, desde el punto de vista estadístico, la mayoría de los alumnos proporciona muestras cuya variabilidad es aceptable en este ítem.

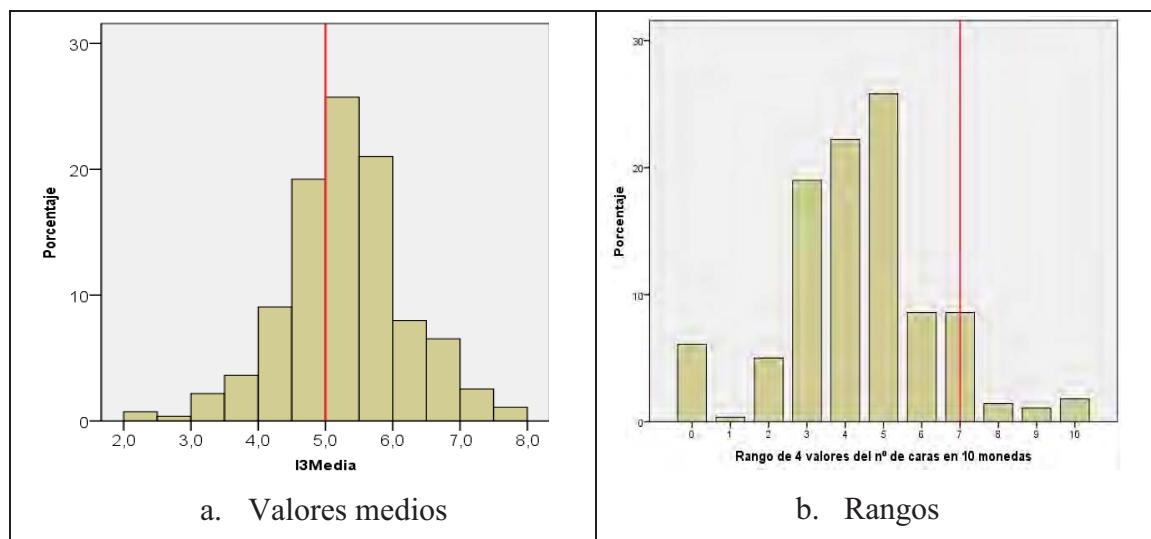


Figura 7. Distribución en el muestro del número medio y rango del nº de caras en 4 lanzamientos de 10 monedas.

Finalmente, el primer análisis conduce a la identificación de una proporción de muestras alta (84%) cuyo valor medio de las muestras es próximo al teórico. Esta observación se complementa con el hecho de que la variabilidad es adecuada en la

mayoría de los casos (52,31%). Por tanto, la proporción de estudiantes que responden a esta cuestión correctamente es mayor que en los dos ítems anteriores. Esto se traduce en que los alumnos presentan una intuición o comprensión adecuada del muestreo cuando el experimento aleatorio está asociado a un número de ensayos menor y los sucesos elementales son equiprobables. No obstante, es conveniente remarcar que a diferencia de los ítems anteriores, en este caso no se proporcionan los valores asociados a un ensayo previo, el cual pueda guiar en la respuesta dada por el alumno.

Comparación por curso

En primer lugar, el análisis de la distribución en el muestreo del valor medio (ver Figura 8a) revela la simetría en ambos grupos, debido a la homogeneidad de frecuencias que se alcanzan en los intervalos simétricos. De hecho, se identifica que son más frecuentes aquellas muestras cuyo valor medio se localiza en el intervalo (4,5, 5,5). En este sentido, los resultados resumidos en la Tabla 5 reflejan un índice de respuestas correctas elevado, donde el 80% de los estudiantes de cada grupo proporcionan muestras cuyo valor medio se sitúa en el intervalo que se ha descrito como correcto.

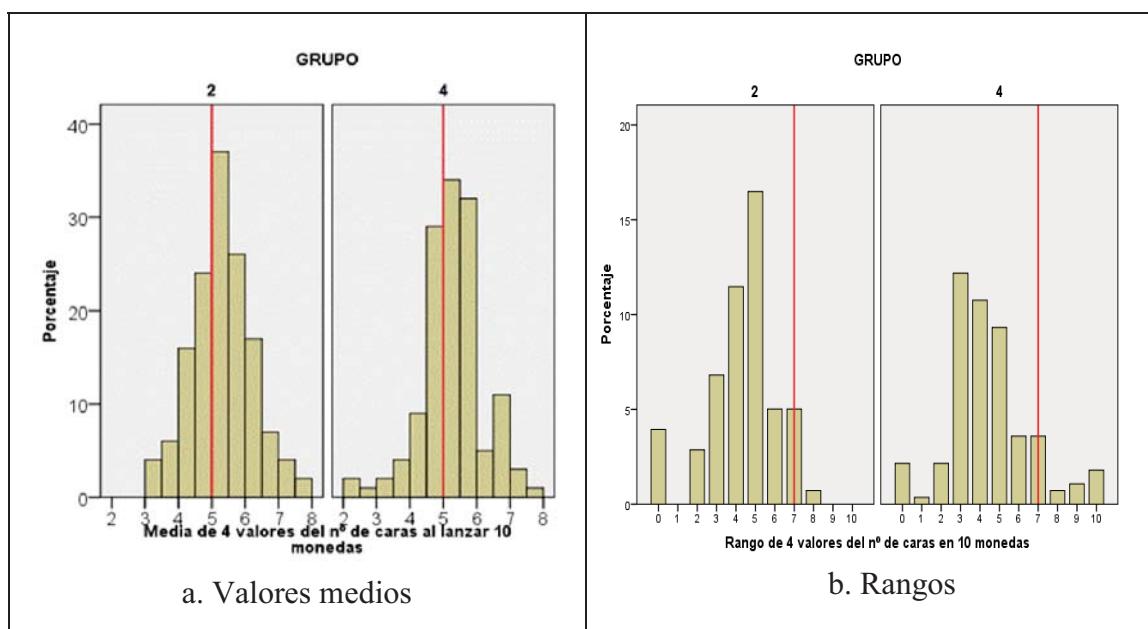


Figura 8. Distribuciones del valor medio y rango del nº de caras en el lanzamiento de 10 monedas por grupo

Por otro lado, se observa que algunos estudiantes presentan el sesgo denominado enfoque del resultado (Konold, 1989) puesto que, presentan una tendencia hacia la

presencia de determinados resultados, que puede deberse a una concepción errónea de la aleatoriedad o del fenómeno que se estudia en particular, como por ejemplo aquellas respuestas constituidas por cuaternas en las que el número de caras es el mismo para cada ensayo (ver Anexo I).

Tabla 5. Frecuencia y porcentaje de alumnos según intervalo en que se sitúa la media de los valores dados

	2ºESO		4ºESO	
	Frecuencia	%	Frecuencia	%
Correcto (3,4-6,6)	131	83	122	84,14
Menor de 3,4	4	2,55	5	3,45
Mayor de 6,6	8	5,10	6	4,13
Otros casos	14	8,90	12	8,30

En segundo lugar, el estudio de la variabilidad recae en el análisis de la distribución del rango (ver Figura 9b) para cada uno de los grupos, que se completa con los datos resumidos en la Tabla 6. En este caso, el análisis comparativo de las frecuencias que se alcanzan en determinados intervalos, revela que las frecuencias siguen una tendencia creciente en el intervalo (3, 5) para el caso del grupo de 2ºESO. Mientras que en el grupo de 4ºESO, se identifica una tendencia decreciente que se traduce en que el grupo de 2ºESO acepta una variabilidad mayor en este experimento.

En conclusión, el análisis de la dispersión de los valores medios (ver Figura 9) revela que el grupo de 2ºESO otorga una mayor variabilidad al estadístico estudiado, la media muestral. Mientras que el grupo 4ºESO no estaría considerando el efecto del tamaño de la muestra y cómo afecta a la variabilidad de las muestras resultantes durante el proceso de muestreo, al proporcionar valores más próximos al valor esperado que conduce a muestras con una concentración mayor de la esperada.

Tabla 6. Frecuencia y porcentaje de alumnos según intervalo en que se sitúa el rango de los valores dados

	2ºESO		4ºESO	
	Frecuencia	%	Frecuencia	%
Hasta 3,2. Alta concentración	35	22,29	47	32,41
3,2-6,4. Adecuada	92	58,60	66	45,52
6,4-9,6. Alta (pero aceptable)	16	10,19	15	10,34
Mayor que 9,6. Excesiva	0	0	5	3,45

Para concluir, discutimos los valores atípicos que se obtienen desde el diagrama de caja para el valor medio asociado a cada grupo (ver Figura 9a). En el caso del grupo

de 2ºESO, se identifica como valores atípicos aquellas respuestas cuyo valor medio se sitúa en los extremos del rango de valores posibles que puede tomar el número de caras. En particular, se identifica la siguiente respuesta: 2S_A125: (7, 6, 8, 10), donde se observa que el alumno proporciona valores muy extremos, mostrando una pobre intuición sobre la frecuencia esperada asociada al número de caras. Por su parte, las respuestas dadas por los alumnos de 4ºESO son: 4AP_A236: (3, 4, 2, 0), 4BS_A272: (5, 1, 3, 0) y 4BS_A284: (9, 8, 7, 6). De manera análoga, se observa que estas muestras se caracterizan por presentar una variabilidad baja, donde el número de caras toma valores muy extremos en relación al rango de valores posibles.

En el caso del estudio del rango, el diagrama de caja (ver Figura 9b) asociado a la muestra de 2ºESO no identifica ningún valor atípico. Además, la mediana señala que el 50% de los alumnos de 2ºESO proporcionan muestras con una dispersión aceptable. Por su parte, el 50% de los estudiantes del grupo de 4ºESO proporciona muestras cuya variabilidad es menor que 4 (valor que toma la mediana representado en la Figura 9b), lo que se traduce en la elaboración de muestras con una variabilidad insuficiente. Por otro lado, el diagrama de caja asociado a este grupo señala la presencia de valores atípicos que se corresponden con muestras cuya variabilidad es excesiva, como ejemplos citamos las siguientes respuestas: 4S_A263: (5, 10, 0, 7) y 4S_A289: (10, 0, 5, 8).

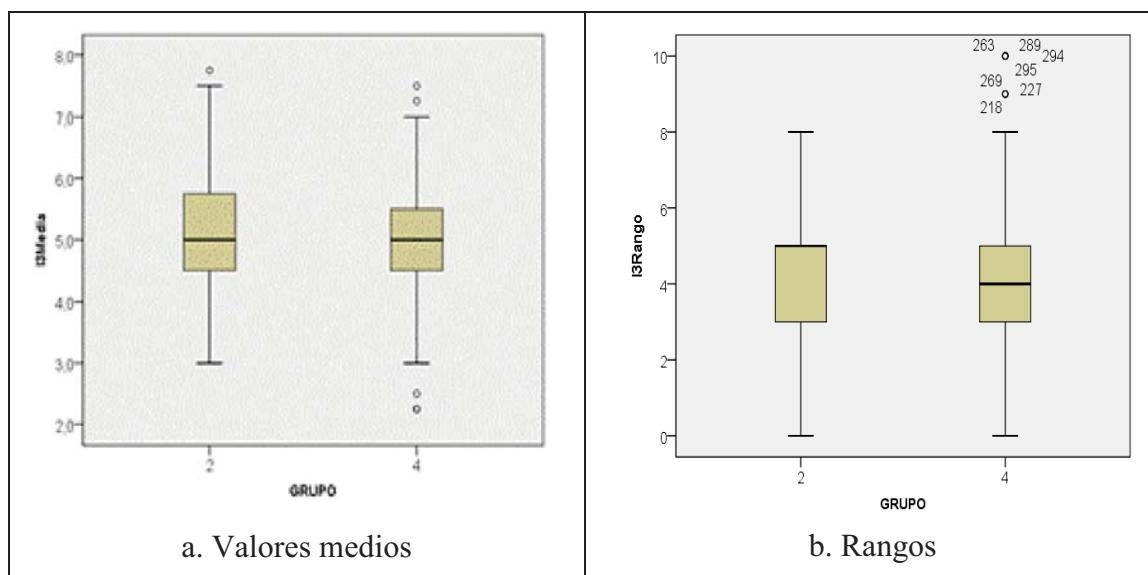


Figura 9. Diagramas de caja del valor medio y rango del nº de caras en el lanzamiento de 10 monedas por grupo

3.5.4. RESULTADOS ÍTEM 4

El análisis de la distribución de las medias refleja que el número de canastas encestadas en un partido donde el jugador lanza 10 veces toma un valor medio de 6,68 y una desviación típica de 1,031. En este caso, la desviación típica asociada al estadístico refleja que la mayoría de los valores medios están muy próximos o son iguales al valor esperado, concentrados en el intervalo (6,5, 8). De hecho, el análisis de las frecuencias de determinados intervalos, refleja que el 74,17% de los estudiantes proporcionan un valor medio localizado en el intervalo definido como respuesta correcta. En este caso, los estudiantes proporcionan, en media, muestras cuya frecuencia refleja de manera adecuada la proporción que guarda en la población.

Por otro lado, la falta de asimetría de los datos (ver Figura 10a) indica que los estudiantes han identificado la no equiprobabilidad del experimento. Es claro que su conocimiento del contexto parece mejor que en el caso de las chinchorras, lo que puede haber favorecido que los estudiantes detecten la no equiprobabilidad del experimento.

En relación con el estudio de la variabilidad, la distribución del rango representada en la Figura 10b alcanza un pico en torno al valor 3. En este caso, los estudiantes presentan una variabilidad adecuada aunque sesgada hacia una concentración alta de los datos. De hecho, se identifica un segundo pico en torno al valor 2, que se asocia con una variabilidad demasiado baja. Por tanto, la tendencia a proporcionar muestras con una variabilidad demasiado baja conduce a determinar que este grupo no identifica el efecto del tamaño de la muestra, y guía sus respuestas en relación al grado de creencia que tienen sobre el fenómeno, presentando el sesgo denominado enfoque de resultado (Konold, 1989).

Por otro lado, un aspecto llamativo, que se identifica en este ítem, es la frecuencia asociada al valor del rango 0, es decir, aquellas respuestas cuya ausencia de variabilidad es extrema, las cuales se corresponden respuestas tales como (7, 7, 7, 7) (ver Anexo I). Este grupo de estudiantes supone el 10,60%. Están haciendo una interpretación determinista del fenómeno, en vez de considerarlo aleatorio. En este caso, los sujetos tienen la creencia de poder contralrar el azar, por lo que este grupo de estudiantes presentan el sesgo que Langer (1982) denomina ilusión de control.

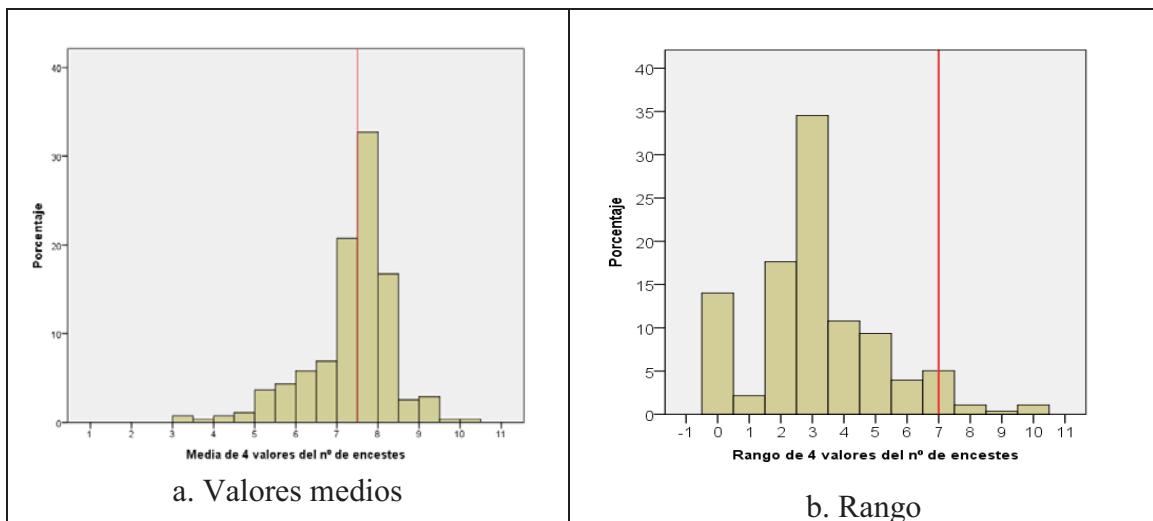


Figura 10. Distribución en el muestro del número medio de encestes en 4 partidos

Comparación por curso

Por un lado, el análisis de la distribución del valor medio para cada uno de los grupos (ver Figura 11a), revela que el valor esperado para el número de encestes se localiza en torno al valor teórico. Concretamente la Tabla 7, que resume la frecuencias con las que aparecen determinados intervalos, refleja que aproximadamente el 70% de los alumnos de cada uno de los grupos proporciona una cuaterna de valores cuyo valor medio se sitúan dentro del intervalo correcto.

Por otro lado, el estudio del la distribución del rango (ver Figura 11b) revela que ambos grupos producen muestras con una dispersión baja. En particular, la Tabla 8 indica que 58,60% del grupo de 2ºESO y el 65,52% del grupo de 4ºESO proporcionan respuestas cuyo rango es menor de 3, que se corresponde con una concentración alta, lo que indica una comprensión parcialmente correcta de la Ley de los Grandes Números, puesto que el estudiante espera que los resultados estén próximos al valor teórico. Esta observación puede deberse a que el alumno guía su razonamiento por su conocimiento previo acerca del fenómeno aleatorio. Es decir, en términos de heurísticas, el resultado obtenido se traduce en la presencia de la heurística de la disponibilidad (Tversky y Kahneman, 1974), que conlleva que el alumno subestime la variabilidad intrínseca a la obtención de muestras de una población dada.

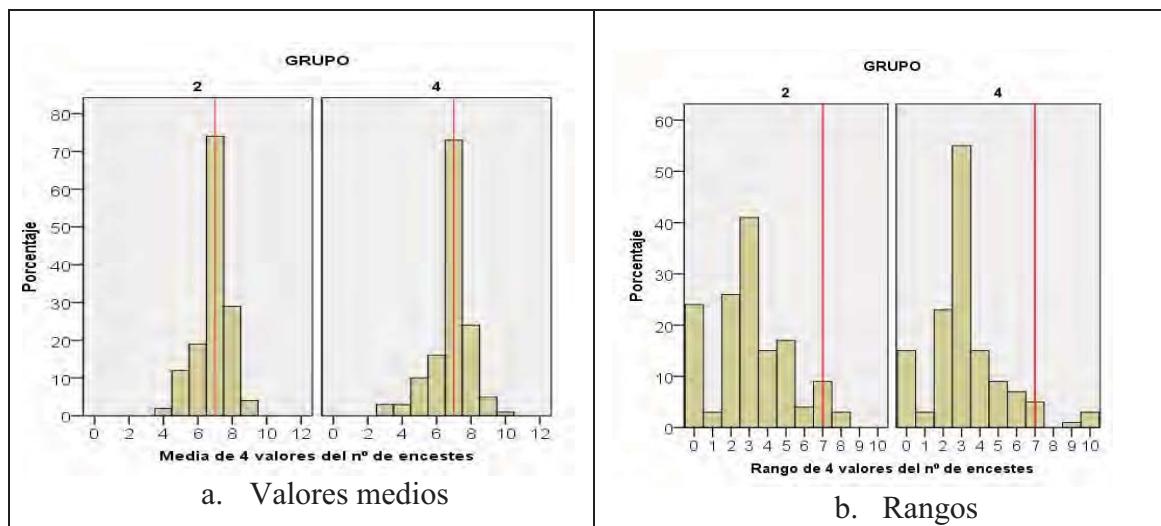


Figura 11. Distribuciones del número medio y rango de 4 valores de nº de encestes

En definitiva, se observa, por un lado, que existe una proporción elevada de estudiantes que producen muestras cuyo valor medio se localiza dentro del intervalo denominado como correcto, por otro lado, el estudio de la variabilidad refleja que los valores presentan una alta concentración. Por tanto, las respuestas de los alumnos revelan que la muestra participante no tiene una comprensión profunda de las dos ideas que caracterizan al concepto de muestreo, es decir, no coordinan los conceptos de representatividad y variabilidad muestral. No obstante, los resultados para este ítem son mejores que los obtenidos para los dos primeros ítems del cuestionario.

Tabla 7. Frecuencia y porcentaje de alumnos según intervalo en que se sitúa la media de los valores dados

	2ºESO		4ºESO	
	Frecuencia	%	Frecuencia	%
Correcto (5,55-8,45)	116	73,89	108	74,48
Menor de 5,55	20	12,10	21	14,48
Mayor de 8,45	4	2,55	6	4,13
Otros casos	17	10,80	10	6,90

Tabla 8. Frecuencia y porcentaje de alumnos según intervalo en que se sitúa el rango de los valores dados

	2ºESO		4ºESO	
	Frecuencia	%	Frecuencia	%
Hasta 3. Alta concentración	92	58,60	95	65,52
3-6. Adecuada	32	20,38	24	16,55
6-9. Alta (pero aceptable)	16	10,19	12	8,28
Mayor que 9. Excesiva	0	0	4	2,76

Finalmente, el análisis del diagrama de caja asociado al valor medio para cada uno de los grupos participantes (ver Figura 12) revela una serie de valores medios atípicos, entre los que destacan los asociados a las siguientes respuestas: 2S_A107: (5, 3, 2, 4), 2P_A48: (5, 4, 3, 2) y 4S_A267: (0, 3, 1, 6). Estos estudiantes proporcionan muestras cuyo valor medio asociada al número de encestes es muy bajo. Por tanto, estas respuestas revelan que los estudiantes no contemplan la información recogida en el enunciado. Por otro lado, se identifican respuestas que reflejan una estimación demasiado elevada del número de encestes: 2P_A30: (8, 9, 10, 10), 4P_A77: (10, 10, 8, 7) y 4S_A268: (10, 9, 9, 10).

En el caso del estudio del rango, se identifican como valores atípicos los asociados a las siguientes respuestas: 2P_A68: (2, 10, 5, 3), 2P_A56: (8, 4, 10, 2) y 2S_A129: (2, 3, 5, 10), 4BP_A227: (10, 2, 1, 0), 4BP_A224: (7, 0, 5, 10), 4BP_A218: (10, 0, 3, 6) y 4BP_A214: (7, 8, 1, 10), donde el rango de dichas respuestas es mayor que 9, por tanto, la variabilidad es muy elevada. No obstante, como se identifica en la Figura 11b la mayoría de los estudiantes de ambos grupos proporcionan respuestas con un rango menor que 8.

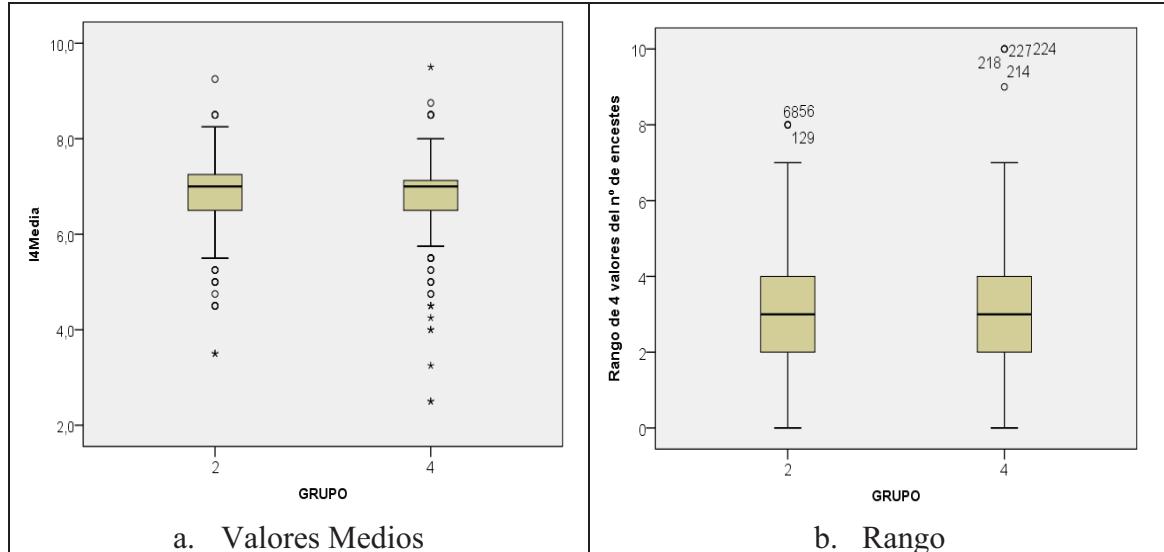


Figura12. Diagramas de caja de la media y rango de 4 valores del nº de encestes

3.6. CONCLUSIONES DEL ANÁLISIS DE LOS DATOS

Este capítulo presenta y describe el instrumento de evaluación diseñado para observar si una muestra de alumnos refleja una adecuada comprensión de los conceptos que se han sido considerados como objeto de análisis para este trabajo. La aplicación

del mismo se lleva a cabo sobre una muestra de 302 estudiantes, de 2ºESO y 4ºESO, de dos institutos de la provincia de Huesca cuyas características son muy similares. La selección de estos dos grupos responde a un interés por observar la existencia de diferencias significativas.

Con respecto a la herramienta de evaluación, se ha considerado alguna de las investigaciones descritas en el Capítulo 2, en particular las investigaciones de Green (1983a) y de Gómez et al. (2014), las cuales, junto con nuestros objetivos iniciales, han guiado en la formulación de los enunciados que constituyen cada ítem. En definitiva, el instrumento de evaluación se corresponde con un cuestionario constituido por cuatro tareas en las que se demanda la generación de muestras. Tras el diseño del cuestionario, el análisis recae en la reflexión de la respuesta esperada para cada ítem que hemos denominado como respuesta correcta o esperada, que nos guía en la determinación de si el sujeto tiene una comprensión adecuada del concepto que evaluamos. Además, la presentación de diferentes experimentos para cada ítem conduce a la presencia de una serie de diferencias que caracterizan a los mismos. Con respecto a esto último, las discrepancias más relevantes de los distintos ítems se corresponden con la equiprobabilidad o no del fenómeno aleatorio y el número de ensayos de la muestra.

Tras la aplicación del instrumento, el interés de este capítulo implica el análisis estadístico de las respuestas elaboradas para cada uno de los ítems mediante técnicas estadísticas. En este sentido, el estudio se fundamenta en la consideración del valor medio asociado a las cuatro muestras que el alumno tiene que elaborar en cada ítem. Así mismo, se contempla el análisis del rango, que se calcula mediante la diferencia entre el valor máximo y el valor mínimo. Este estadístico nos permite inferir aspectos acerca de la variabilidad otorgada a las muestras. Posteriormente, se lleva a cabo un análisis similar para cada uno de los grupos participantes con el objetivo de identificar si existen posibles diferencias entre los dos grupos.

3.6.1. SÍNTESIS DE RESULTADOS

A continuación, se lleva a cabo una síntesis de los resultados del cuestionario para cada uno de los ítems. Para ello, presentamos en la Tabla 3.6.1. tanto las medias y como las desviaciones típicas de los resultados teóricamente esperados en cada ítem, así como las de los valores observados en el estudio, tanto globalmente, como por cursos.

Desde el análisis de los enunciados, se identifica que el ítem 1 y 4 hacen referencia a fenómenos aleatorios cuyos sucesos elementales no son equiprobables a diferencia de las otras dos cuestiones. Por su parte, el primer y segundo ítems hacen referencia a un experimento cuyo número de ensayos para cada muestra toma un valor mayor que en los otros dos. En definitiva, las dos características citadas nos permiten observar si el sujeto identifica la equiprobabilidad o no del fenómeno, así como la representatividad y variabilidad muestral. En relación con este último aspecto, la diferencia en el número de ensayos permite analizar el efecto del tamaño de la muestra en relación a la variabilidad aceptada para la misma.

Tabla 3.6.1. Síntesis de resultados

Item	Teórico		Global		2º curso		4º curso	
	Valor esperado	D. Típica						
1 B (100,0,68)	68	4,7	57,9	12,8	56,8	13,7	61,1	8,7
2 B (100, 0,5)	50	5	51,2	7,6	51,9	8,8	50,6	6,4
3 B (10, 0,5)	5	1,6	5,2	0,9	5,1	0,8	5,1	0,9
4 B (10, 0,7)	7	1,45	6,7	1	6,8	0,1	6,6	0,1

Globalmente el estudio de la Tabla 3.6.1. nos muestra que la estimación del valor esperado es buena, en general, excepto en el primer ítem, tanto globalmente, como por grupo. Pensamos que los estudiantes tienen una buena comprensión o intuición del valor esperado en el muestreo y el fallo en el primer ítem es debido a la falta de experiencia con el experimento que consiste en lanzar las chinchorras. Por el contrario, el resto de los enunciados, en los que se presentan fenómenos aleatorios como lanzar una moneda o jugar al baloncesto, se corresponden con contextos familiares para el alumno.

No obstante el aspecto más relevante, es la falta de comprensión de la variabilidad muestral por las diferencias que hay entre las desviaciones típicas teóricas y las observadas. Además, coincidiendo con la investigación de Gómez et al. (2014), observamos que no existe una intuición simultánea de la convergencia al valor esperado y de la variabilidad muestral. Por tanto, creemos que este es un punto en el que se requiere mejorar la enseñanza.

3.6.2. COMPARACIÓN DE RESULTADOS SEGÚN EL TAMAÑO DE LA MUESTRA

Como se ha citado en las líneas anteriores, el análisis recae en comparar los resultados obtenidos en los ítems 1 y 2 con los ítems 3 y 4. En este sentido, en las dos primeras tareas se identifica que los estudiantes independiente del curso proporcionan muestras con una variabilidad extrema, siendo la frecuencia mayor en el grupo de 2ºESO. Esta observación se desprende, a su vez, de los datos resumidos en la Tabla 3.6.1., especialmente para el primer ítem. En general, el análisis de cada ítem revela que para los dos primeros ítems la mitad de los alumnos de 2ºESO proporcionan muestras con una variabilidad extrema, mientras que este porcentaje es del 40% en el caso de la muestra de 4ºESO. Por tanto, se observa que los estudiantes no presentan una adecuada comprensión de la variabilidad cuando el tamaño de la muestra es mayor.

Por su parte, las muestras elaboradas para los dos últimos ítems presentan una variabilidad menor que la esperada con pocas diferencias en ambos cursos.

En resumen, estos resultados reflejan que los estudiantes conceden a las grandes muestras más variabilidad que la teórica y menos variabilidad que la teórica a las muestras pequeñas. Es posible que ello encubra, en algunos estudiantes tanto la heurística de la representatividad (en su estimación en las muestras pequeñas) como la representatividad (en su respuesta a las muestras grandes).

Por otro lado, aunque la mayoría de las respuestas elaboradas proporcionan valores cuya media se localiza dentro del intervalo definido como correcto, el apunte anterior conlleva a determinar que no existe una adecuada coordinación de las dos ideas fundamentales para una buena comprensión del concepto de muestreo: representatividad muestral y variabilidad muestral. Es decir, el estudiante proporciona, por ejemplo, un conjunto de cuatro muestras cuyo valor medio para el número de chinchetas que caen con la punta hacia arriba es correcto, sin embargo, la variabilidad de dichas muestras es excesiva. Por tanto, el estudiante no presenta una intuición adecuada de la Ley de los Grandes Números.

En definitiva, los estudiantes no comprenden el efecto del tamaño de la muestra sobre la variabilidad en el muestreo, resultado que aparece en la investigación de Serrano (1996), Shaughnessy et al. (2004), Gómez (2014) y Gómez et al. (2014).

Además, los resultados obtenidos se sitúan en la línea de las observaciones o resultados aportados por las investigaciones revisadas en la literatura, puesto que ponen en relieve la dificultad que los alumnos presentan en la comprensión de la variabilidad intrínseca al proceso de muestreo.

3.6.3. COMPARACIÓN DE RESULTADOS EN EXPERIMENTOS CON SUCESOS EQUIPROBABLES O NO

Con el objetivo de examinar si los estudiantes son capaces de identificar la equiprobabilidad o no del fenómeno aleatorio se han diseñado cuestiones que permitan analizar si se identifica la asimetría del fenómeno presentado. En este caso, los fenómenos presentados en el primer y último ítem se corresponden con fenómenos que no cumplen el principio de equiprobabilidad. Desde el análisis de los datos para el segundo y tercer ítem, se identifica que la muestra de participantes percibe la equiprobabilidad del fenómeno aleatorio, siendo este último el mismo en ambas cuestiones. Por tanto, las siguientes líneas suponen una síntesis de este aspecto para las otras dos cuestiones.

Por un lado, el análisis de la distribución del valor medio de los datos recogidos para el primer ítem revela que en ambos grupos se identifica una proporción de estudiantes que no identifican la asimetría del dispositivo, debido al sesgo de equiprobabilidad o a la heurística de la representatividad, resultado que aparece en los trabajos de Gómez (2014), Gómez et al. (2014) y Serrano (1996). Esta heurística se corresponde con la tendencia de compensar el conjunto de respuestas mediante la consideración de valores donde el número de chinches que caen con la punta hacia arriba es menor, por tanto, no se considera el resultado del profesor.

En el caso del último ítem, el análisis de la distribución del valor medio en el muestreo refleja que los estudiantes identifican la no equiprobabilidad del experimento. En este caso, creemos que el conocimiento o la familiaridad del contexto favorecen la comprensión del fenómeno aleatorio. No obstante, la escasa variabilidad asociada a las muestras indica que algunos estudiantes dan al fenómeno una interpretación prácticamente determinista.

CAPÍTULO 4.

CONCLUSIONES

4.1. INTRODUCCIÓN

Este último capítulo implica una reflexión sobre los objetivos y las hipótesis iniciales en relación con los resultados obtenidos del estudio realizado. Así mismo, se reflexiona sobre las limitaciones de este trabajo, y las posibles vías de investigación que quedan abiertas. Finalmente, el último punto está dirigido a analizar las implicaciones que el estudio tiene para la enseñanza.

4.2. CONCLUSIONES RESPECTO A LOS OBJETIVOS

Este trabajo tiene como finalidad analizar y evaluar la comprensión que un grupo de estudiantes de la etapa educativa obligatoria de secundaria presenta sobre algunas ideas asociadas al concepto de muestreo. Desde este objetivo general se definen un conjunto de objetivos secundarios que enmarcan el estudio. Por tanto, las siguientes líneas recogen una breve reflexión de los mismos, tras la realización de la parte empírica del trabajo:

Objetivo 1. Identificar, describir y comparar los contenidos matemáticos presentados en los documentos curriculares.

Para cumplir este objetivo se hizo una revisión de los documentos curriculares. A partir de ella se identifica que la LOMCE busca un progreso gradual de la enseñanza y aprendizaje de los contenidos del bloque de estadística y probabilidad, de manera análoga a la anterior ley de educación, la LOE. No obstante, la diferencia radica en que este nuevo currículo adelanta la enseñanza de determinados contenidos matemáticos. En relación con esta observación, la enseñanza de la probabilidad desde el significado clásico se inicia en el primer curso de la etapa de secundaria, mientras que la LOE reserva este contenido para el tercer curso de la misma.

Objetivo 2. Realizar una síntesis de las investigaciones previas.

Con la finalidad de cumplir este objetivo, se lleva a cabo una revisión de la literatura existente, la cual nos guía en nuestro trabajo de investigación. Este último

aspecto se corresponde con el hecho de que el instrumento de evaluación contempla alguna de las cuestiones que han formado parte de las investigaciones anteriores consideradas. Además, esta característica permite contrastar los resultados obtenidos con los de los trabajos previos. En este sentido, el análisis de los datos revela que los estudiantes siguen sin presentar una adecuada comprensión de la variabilidad muestral, tal y como se detecta en aquellas investigaciones que contemplan como finalidad la evaluación de conceptos asociados al muestreo.

Objetivo 3. Analizar la comprensión de la relación entre el valor de una proporción en la población y la frecuencia relativa esperada en muestras tomadas de dicha población en alumnos de Educación Secundaria Obligatoria.

Objetivo 4. Evaluar y analizar la comprensión de la variabilidad en el muestro y el efecto del tamaño de la muestra sobre dicha variabilidad en alumnos de Educación Secundaria Obligatoria.

La consecución de estos dos objetivos se corresponde con la aplicación del cuestionario, que hemos diseñado, sobre una muestra amplia de estudiantes de 2º y 4º curso de Educación Secundaria Obligatoria. El estudio de las respuestas elaboradas por los participantes se describe en el Capítulo 3, donde analizamos con detalle los resultados de cada ítem. En primer lugar, realizamos un análisis sobre la muestra total, que se completa con el análisis comparativo según el curso. Posteriormente, el interés recae en analizar las respuestas según las diferencias debidas al tamaño de la muestra, así como de si el fenómeno aleatorio descrito cumple o no el principio de equiprobabilidad.

Todo ello nos proporciona información relevante para identificar si los estudiantes presentan una comprensión adecuada de las ideas caracterizadas en los objetivos anteriores. Las conclusiones asociadas a los mismos se discuten de manera más detallada al reflexionar sobre la hipótesis 2.

4.3. CONCLUSIONES RESPECTO A LAS HIPÓTESIS INICIALES

Desde el interés por analizar la comprensión que un grupo de estudiantes presenta sobre diferentes ideas asociadas al concepto de muestreo, surgen una serie de hipótesis que se describen más detalladamente durante el planteamiento del problema (Capítulo 1). Tras la aplicación del cuestionario y el análisis de los datos obtenidos, es relevante

reflexionar sobre los resultados obtenidos en relación con las hipótesis descritas previamente.

H1. Los contenidos curriculares tanto del Decreto de Enseñanzas Mínima como del Curículo básico contemplan o reflejan los aspectos esenciales asociados al concepto de muestreo que son objeto de investigación.

En este caso, el análisis de los documentos curriculares, tanto a nivel estatal como autonómico, muestra que ambos textos consideran la introducción gradual de las ideas necesarias para la comprensión adecuada y profunda del concepto. Además, se remarca la importancia de que el estudiante muestre una actitud crítica tanto en la elaboración como en la interpretación de estudios estadísticos.

H2. Algunos alumnos presentan una comprensión insuficiente sobre las ideas básicas asociadas al concepto de muestreo.

Esta hipótesis se cumple solo en parte, pues se obtiene una gran proporción de respuestas correctas en la estimación de la proporción en los problemas planteados. El análisis de los datos, revela también que la estimación es mejor, cuando el fenómeno aleatorio presentado cumple la condición de equiprobabilidad, independiente del tamaño de la muestra. No obstante, el análisis de los datos asociados al ítem 1 revela que la muestra de estudiantes presenta tanto el sesgo de equiprobabilidad como la heurística de representatividad. Esta última se identifica en aquellos casos en los que el alumno trata de compensar los resultados, proporcionando muestras en los que el número de chinchetas que caen con la punta hacia abajo es mayor. Esta heurística conlleva a la estimación errónea de la frecuencia. Aunque estos resultados mejoran considerablemente en la última de las cuestiones, debido a la familiaridad del estudiante con la situación problemática planteada.

Por otro lado, el estudio realizado refleja que la mayoría de los alumnos tiene una comprensión insuficiente de la variabilidad intrínseca al proceso de muestreo, que se acentúa en aquellas muestras cuyo tamaño es mayor. En estos ítems, la mayoría de las respuestas presentan muestras cuyos valores quedan caracterizados por una variabilidad excesiva. Por tanto, se confirma una mayor dificultad en la idea de variabilidad. También encontramos la presencia de respuestas en la que se identifica una ausencia de la misma (ver Anexo 1).

H3. El análisis de las respuestas del instrumento de evaluación refleja un progreso gradual en los contenidos evaluados en los alumnos del 4º curso.

La formulación de esta hipótesis responde a la configuración de la muestra participante, la cual está formada por alumnos que actualmente están cursando 2º ESO o 4º ESO. A pesar de que se esperaba que los alumnos de 4º ESO obtuvieran mejores resultados, debido a factores tales como la madurez y su conocimiento, el análisis de las respuestas revela que las diferencias entre los dos grupos escolares no son muy altas, aunque el grupo de 4º ESO suele obtener mejores resultados.

En síntesis, la reflexión sobre las diferentes hipótesis pone en relieve la desconexión existente entre los contenidos descritos en los documentos curriculares y el conocimiento que los estudiantes muestran sobre dichos contenidos matemáticos.

4.4. LIMITACIONES DEL TRABAJO

Las limitaciones que se identifican tienen su respuesta en la naturaleza del mismo. En este sentido, el trabajo de investigación es un estudio exploratorio cuya finalidad es identificar la comprensión que los alumnos presentan acerca de diversos conceptos asociados al estudio del muestreo. Para lo cual, el instrumento que se aplica consta de cuatro tareas de generación de muestras de un tamaño determinado sobre el fenómeno aleatorio que se describe. En primer lugar, los ítems no exigen que el estudiante razonne acerca de los valores que redacta. Aunque la elección de este tipo de ítems responde a motivos relacionados con la dificultad del cuestionario, es claro que este aspecto limita la investigación, porque no se puede analizar todo el proceso cognitivo que conduce a la respuesta.

Por otro lado, los alumnos de 2º ESO no habían recibido instrucción sobre contenidos referentes al bloque de probabilidad y estadística. En el caso de los estudiantes de 4º ESO este factor es mucho más complejo, puesto que algunos de los alumnos habían visto contenidos propios de la estadística durante el curso anterior. Otros, sin embargo, habían trabajado la probabilidad desde el punto de vista clásico. Finalmente, un tercer grupo, la mayoría de los casos, expresa no haber visto contenidos propios del bloque de estadística y probabilidad durante los cursos anteriores. La dificultad de controlar este factor se corresponde el hecho de que los alumnos no recordaban que contenidos habían visto y mostraban su falta de conocimiento sobre los

mismos, a pesar de haberlos trabajado. Aunque esta variable puede enmascarar los resultados obtenidos en el cuestionario, el objetivo de nuestro estudio es observar la comprensión o intuición que los estudiantes presentan sobre conceptos que no han sido trabajados de manera explícita en el aula, y que por tanto no han desarrollado un aprendizaje profundo de los mismos en ninguno de los casos. Por tanto, este factor señalado, aunque es relevante, no afecta de manera sustancial en el proceso de la tarea investigadora.

4.5. FUTURAS LÍNEAS DE INVESTIGACIÓN

En primer lugar, los resultados revelan una falta de comprensión adecuada sobre el concepto de muestreo, siendo más acusada para el caso de la variabilidad muestral. Este hecho conduce a que el estudiante presente dificultades para entender las ideas fundamentales asociadas a la inferencia estadística, entre las que se sitúa la idea de muestreo. Entonces, desde la consideración de las limitaciones citadas en el apartado anterior, consideramos que en el futuro se podría plantear una investigación más exhaustiva. Por ello, se nos plantea la necesidad de reflexionar sobre una investigación en la que se diseñara un proceso de instrucción adaptado para los dos cursos considerados en este trabajo, con la finalidad de observar si existe una mejora en la comprensión de las ideas tras la enseñanza de las mismas.

Por otro lado, se puede llevar a cabo una réplica del estudio donde se amplíe la muestra participante, de modo que se considere un diseño transversal en el que se participen estudiantes de diferentes niveles educativos.

Finalmente, las limitaciones debidas al tipo de ítems considerados implica el planteamiento de una investigación en la que se modifique los enunciados. En este sentido, se podría demandar al alumno que describiera, además, el razonamiento que guía en su respuesta. Esta modificación puede ser viable si se considera la elaboración de un material que favorezca el proceso de aprendizaje de dichos conceptos, permitiendo al alumno la consecución de herramientas que le ayuden en la expresión verbal de la tarea.

Para concluir, la importancia que adquieren las ideas de muestreo y distribución muestral para adquirir una adecuada comprensión de la estadística (Ben-Zvi et al., 2015), justifica la necesidad de llevar a cabo investigaciones que favorezcan el

aprendizaje de estos conceptos desde los primeros cursos de la secundaria, con la finalidad de no dejar este contenido a los últimos cursos del Bachillerato, y principalmente, a la etapa universitaria. Por otro lado, Watson y Moritz (2000) señalan que la investigación sobre el desarrollo cognitivo del concepto de muestreo tiene una presencia escasa en la literatura.

4.6. IMPLICACIONES PARA LA ENSEÑANZA

Finalmente, cabe destacar las implicaciones que los resultados de este trabajo tienen sobre la enseñanza. En primer lugar, los resultados revelan que los estudiantes presentan heurísticas que guían sus reflexiones dando lugar a la presencia de sesgos en las respuestas aportadas que conducen a una respuesta incorrecta de la tarea. Además, como los resultados obtenidos coinciden con las investigaciones previas, esta observación pone en relieve 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 desarrollos de manera gradual los contenidos trabajados en la etapa anterior.

Por otro lado, desde el estudio realizado, se identifica la falta de protagonismo que adquiere la enseñanza de la Probabilidad desde el enfoque frecuencial, a pesar de que las directrices curriculares sostienen un cambio en el enfoque desde el que iniciar la enseñanza de la probabilidad, considerándose su significado frecuencial. En relación con nuestro problema, señalamos que esta perspectiva de la Probabilidad supone un puente entre la Probabilidad y la Estadística, por lo que es imperante que el docente adopte una actitud crítica acerca de las decisiones que lleva a cabo en el aula para, así como valore la inclusión de un determinado contenido en el aula. Con la finalidad de beneficiar el aprendizaje del alumno sobre los contenidos matemáticos, y los citados en esta memoria, en particular.

ANEXO 1. ESTUDIO DE CASOS EXTREMOS

Las siguientes tablas reflejan aquellos estudiantes que responden a la tarea con una cuaterna constituida por valores con dispersiones extremas caracterizadas por la ausencia de variabilidad.

ÍTEM 1

Caso extremo. Ausencia de variabilidad	2ºESO		4ºESO	
	Frec.	% (respecto al total de alumnos de 2ºESO que realiza la prueba)	Frec.	% (respecto al total de alumnos de 4ºESO que realiza la prueba)
(50, 50, 50, 50)	7	4,46	5	3,45
(68, 68, 68, 68)	6	3,82	1	0,69
(70, 70, 70, 70)	1	0,64	2	1,38
(80, 80, 80, 80)				
(100, 100, 100, 100)	2	1,27		
Total	16	10,19	8	5,52

ÍTEM 2

Caso extremo. Ausencia de variabilidad	2ºESO		4ºESO	
	Frec.	% (respecto al total de alumnos de 2ºESO que realiza la prueba)	Frec.	% (respecto al total de alumnos de 4ºESO que realiza la prueba)
(50, 50, 50, 50)	7	4,46	8	5,52
(53, 53, 53, 53)	4	2,55	1	0,68
(47, 47, 47, 47)	1	0,64		
Total	12	7,64	9	6,21

ÍTEM 3

Caso extremo. Ausencia de variabilidad	2ºESO		4ºESO	
	Frec.	% (respecto al total de alumnos de 2ºESO que realiza la prueba)	Frec.	% (respecto al total de alumnos de 4ºESO que realiza la prueba)
(5, 5, 5, 5)	7	4,46	6	4,14

ÍTEM 4

Caso extremo. Ausencia de variabilidad	2ºESO		4ºESO	
	Frec.	% (respecto al total de alumnos de 2ºESO que realiza la prueba)	Frec.	% (respecto al total de alumnos de 4ºESO que realiza la prueba)
(7, 7, 7, 7)	19	12,10	13	8,97
(5, 5, 5, 5), (6, 6, 6, 6), (8, 8, 8, 8)	3	1,91	1	0.68
Total	22	14,01	14	9,66

REFERENCIAS

- Arteaga, P., Batanero, C., Cañadas, G. y Contreras, J. M. (2011). Las tablas y gráficos estadísticos como objetos culturales, *Números* 76, 55-67.
- Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools*. Utrecht, the Netherlands: Beta Press.
- Ball, D. L., Lubienski, S. T. y Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. En V. Richardson (Ed.), *Handbook of research on teaching* (pp. 433-456). Washington, DC: American Educational Research Association.
- Batanero, C. (2004). Ideas estocásticas fundamentales. ¿Qué contenidos se debe enseñar en la clase de probabilidad? En J. A. Fernández (Ed.), *Actas do I Encontro de Probabilidades e Estatística na Escola* (pp. 9-30). Braga: Portugal: Universidad de Minho.
- Batanero, C. (2013). Del análisis de datos a la inferencia: Reflexiones sobre la formación del razonamiento estadístico. *Cuadernos de Investigación y Formación en Educación Matemática*, 8(11), 277-291.
- Batanero, C. y Díaz, C. (2015). Aproximación informal al contraste de hipótesis. En J. M. Contreras, C. Batanero, J. D. Godino, G.R. Cañadas, P. Arteaga, E. Molina, M.M. Gea y M.M. López (Eds.), *Didáctica de la Estadística, Probabilidad y Combinatoria*, 2, 135-144.
- Batanero, C., Díaz, C., Contreras, J. M. y Roa, R. (2013). El sentido estadístico y su desarrollo. *Números* 83, 7-18.
- Batanero, C., Gea, M., Arteaga, P. y Contreras, J.M. (2014). La estadística en la educación obligatoria: Análisis del currículo español. *Revista Digital Matemática, Educación e Internet* 14(2). Recuperado de: <http://revistas.tec.ac.cr/index.php/matematica/article/view/1663>
- Batanero, C. y Godino, J. (2005). Perspectivas de la educación estadística como área de investigación. En L. Blanco (Ed.), *Líneas de investigación en Didáctica de las Matemáticas* (pp. 203-226). Badajoz: Universidad de Extremadura.
- Batanero, C., Godino, J. D., Green, D. R., Holmes, P. y Vallecillos, A. (1994). Errores y dificultades en la comprensión de los conceptos estadísticos elementales. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.

- Ben-Zvi, D., Bakker, A. y Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, 88(3), 291-303.
- Bisquerra, R. (1989). *Métodos de investigación educativa*. Barcelona: P.P.U.
- 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 – A joint ICMI/IASE study* (pp. 57-69). Dordrecht: Springer.
- Cañizares, M. J. (1997). *Influencia del razonamiento proporcional y combinatorio y de creencias subjetivas en las intuiciones probabilísticas primarias*. Tesis doctoral. Universidad de Granada.
- Chance, B., delMas, R. C. y Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295-323). Amsterdam: Kluwer.
- delMas, R., Garfield, J., & Chance, B. (2004). Using assessment to study the development of students' reasoning about sampling distributions. Paper presented at the Annual Meeting of the American Educational Research Association, April 13, San Diego, CA. [Online: www.gen.umn.edu/faculty_staff/delmas/AERA_2004_samp_dist.pdf]
- Departamento de Educación, Cultura y Deporte (2007). *Orden de 9 de mayo de 2007, por el que se desarrolla el currículo de la Educación Secundaria Obligatoria y se autoriza su aplicación en los centros docentes de la Comunidad de Aragón*.
- Díaz, C. (2003). Heurísticas y sesgos en el razonamiento probabilístico. Implicaciones para la enseñanza de la estadística. *Actas del 27 Congreso Nacional de Estadística e Investigación Operativa*. Lérida: Sociedad de Estadística e Investigación Operativa. [CD- ROM].
- Falk, R. y Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, 104, 301-318.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Garfield, J. y Gal, I. (1999). Teaching and assessing statistical reasoning. In L. Stiff (Ed.), *Developing mathematical reasoning in grades K-12* (pp. 207-219). Reston, VA: National Council Teachers of Mathematics.

- Godino, J. D. (1996). Mathematical concepts, their meanings and understanding. En L. Puig y A. Gutiérrez (Eds.), *Proceedings of the 20th PME Conference* (v.2, 417-424). Universidad de Valencia.
- Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques* 22 (2 y 3), 237-284.
- Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, 14(3), 325-355.
- Godino, J. D., Batanero, C. y Cañizares, M. J. (1987). *Azar y probabilidad. Fundamentos didácticos y propuestas curriculares*. Madrid: Síntesis.
- Godino, J. D. Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education*, 39 (1-2), 127-135.
- Gómez, E. (2014). *Evaluación y desarrollo del conocimiento matemático para enseñar la probabilidad en futuros profesores de educación primaria*. Tesis Doctoral. Universidad de Granada.
- 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.
- Green, D. R. (1983a). 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.
- Green, D. R. (1983b). From thumbtacks to inference. *School Science and Mathematics*, 83(7), 541-551.
- Green, D. R. (1991). A longitudinal study of children's probability concepts. En D. Vere Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 320-328). Dunedin: Universidad de Otago.
- 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). Springer.
- Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics* 6, 187-205.
- Holmes, P. y cols. (1980). *Statistics in your world*. (Slough: Foulsham Educational).

- Jones, G. A. (Ed.). (2006). Exploring probability in school: Challenges for teaching and learning (Vol. 40). Springer Science & Business Media.
- 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. Recuperado de: https://www.stat.auckland.ac.nz/~iase/publications/rt08/T4P9_Kadijevich.pdf
- Kahneman, D., Slovic, P. y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, 6, 59-98.
- Langer, E. J. (1982). The illusion of control. En D. Kahneman, P. Slovic, y A. Tversky (Eds.), *Judgment under uncertainty: Heuristic and biases* (pp. 231-238). New York: Cambridge University Press.
- Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics*, 23, 557-568.
- Liu, Y. y Thompson, P. W. (2009). Mathematics teachers' understandings of proto-hypothesis testing. *Pedagogies*, 4 (2), 126-138.
- MEC (2007). *Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria*. España: Ministerio de Educación y Cultura.
- MEC (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*.
- Moore, D. S. (1991). Teaching statistics as a respectable subject. En F. Gordon y S. Gordon (Eds.), *Statistics for the twenty-first century* (pp.14-25). Mathematical Association of America.
- Saldanha. L. y Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257-270.
- Schuyten, G. (1991). Statistical thinking in psychology and education. En D. Vere-Jones (Ed.) *Proceeding of the Third International Conference on Teaching Statistics* (pp. 486-490). Otago, Nueva Zelanda: International Statistical Institute),

- 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.
- Shaughnessy, J.M., Ciancetta, M. y Canada, D. (2004). Types of student reasoning on sampling tasks. En M.J. Høines y A.B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol.4, pp. 177-184). Bergen, Noruega: International Group for the Psychology of Mathematics Education.
- Tversky, A. y Kahneman, D. (1974): Judgement under uncertainty: Heuristics and biases. *Science.185*, 1124-1131.
- 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.
- Vanhoof, S., Castro Sotos, A. E., Onghena, P. y Verschaffel, L. (2007). Students' reasoning about sampling distributions before and after the sampling distribution activity. *Proceedings of the 56th Session of the International Statistical Institute*. Lisboa: International Statistical Institute. Recuperado de: http://iase-web.org/documents/papers/isi56/CPM80_Vanhoof.pdf.
- Watson, J. M. (2004). Developing reasoning about samples. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 277–294). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Watson, J., Beswick, K., y Brown, N. (2012). *Educational research and professional learning in changing times: The MARBLE experience*. Springer Science & Business Media.
- Watson, J. M. y Moritz, J. B. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education*, 31(1) 44-70.
- Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67 (3), 221-248.