revista
PARANAENSE
DE EDUCAÇÃO
MATEMÁTICA

LA INFERENCIA ESTADÍSTICA EN LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD EN ANDALUCÍA

M. del Mar López-Martín*
Carmen Batanero**
Carmen Díaz-Batanero***
M. Magdalena Gea****

Resumen: El objetivo de este estudio ha sido analizar el contenido matemático de los problemas de inferencia propuestos desde 2003 a 2014 en las Pruebas de Acceso a la Universidad del Distrito Andaluz en la especialidad de Bachillerato de Ciencias Sociales. Mediante un análisis de contenido se ha identificado el campo de problema que corresponde a cada ejercicio. Se utiliza un ejemplo en cada categoría para identificar el contenido matemático evaluado, analizando la diferencia entre campos de problemas y la correspondencia con los estándares de aprendizaje evaluables propuestos en el nuevo currículo. Finalmente se estudia la distribución global y por año de los diferentes campos de problemas. Los resultados obtenidos del análisis pueden servir para la elaboración de pruebas futuras y preparar a los estudiantes que tienen que enfrentarse a las mismas.
Palabras-clave: Inferencia estadística. Pruebas de Acceso a la Universidad. Análisis de problemas. Evaluación. Bachillerato de Ciencias Sociales.

STATISTICAL INFERENCE PROBLEMS IN THE UNIVERSITY ENTRANCE TEST IN ANDALUCÍA

Abstract: The aim of this study was to analyse the content of inference problems proposed from 2003-2014 for the university entrance tests in Andalucía (Spain) in the specialty of Social Sciences. Through a content analysis we identified the problem field in each exercise. The analysis of one example in each category reveals differences in the assessed content and the correspondence with the learning standards proposed in the new curriculum. We finally study the distribution of problem fields by year and globally. The results can be used to improve future tests and prepare students who have to face them.
Keywords: Statistical inference. University entrance test. Problem analysis. Assessment. Social Sciences High School.

Introducción

La inferencia estadística ha adquirido gran importancia en las últimas décadas, generalizándose su enseñanza en la mayoría de los grados y másteres universitarios en las universidades españolas. Los alumnos tienen la posibilidad de introducirse en el tema cursando los contenidos del Bloque de Estadística y Probabilidad (MEC, 2007) incluidos en la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II, en el Bachillerato de

Humanidades y Ciencias Sociales. Los contenidos fijados son los siguientes (p.45476): Implicaciones prácticas de los teoremas: Central del límite, de aproximación de la Binomial a la Normal y Ley de los Grandes Números. Problemas relacionados con la elección de las muestras. Condiciones de representatividad. Parámetros de una población. Distribuciones de probabilidad de las medias y proporciones muestrales. Intervalo de confianza para el parámetro p de una distribución binomial y para la media de una distribución normal de desviación típica conocida. Contraste de hipótesis para la proporción de una distribución binomial y para la media o diferencias de medias de distribuciones normales con desviación típica conocida.

Observamos que el contenido pretendido abarca prácticamente un curso elemental de inferencia estadística a nivel universitario, incluso cuando esta asignatura está dirigida a estudiantes que pudieran tener pocos conocimientos de álgebra y de cálculo que les permitan realizar un estudio formal de la inferencia. Estos contenidos se mantienen con pocos cambios en el nuevo Decreto de Educación Secundaria y Bachillerato (MECD, 2015). Es posible, por tanto, que los estudiantes no tengan tiempo suficiente para estudiar y asimilar todo este contenido, al que hay que añadir los correspondientes de probabilidad (también bastante completos), así como los correspondientes al resto de los bloques temáticos.

La importancia que dentro del aula se da a cada tema recogido en las orientaciones curriculares, en particular, para los estudiantes de segundo de Bachillerato viene implícitamente determinada por las Pruebas de Evaluación del Distrito Andaluz de Acceso a la Universidad (en adelante PAU). Aunque la finalidad principal de las pruebas PAU es evaluar los conocimientos y capacidades adquiridas por los futuros universitarios durante su etapa de Bachillerato (primer y segundo curso), también se utilizan actualmente para seleccionar a los estudiantes que quieren ingresar en titulaciones y centros de estudios universitarios determinados.

Para analizar si los contenidos de inferencia de las pruebas están directamente relacionados con los recogidos en el currículum de Bachillerato para los estudiantes de Ciencias Sociales, así como el peso relativo que reciben dichos contenidos en el total de la prueba, en este trabajo analizamos los problemas de inferencia propuestos en las pruebas PAU de la Comunidad Autónoma de Andalucía celebradas desde 2003 hasta 2014. Dado que el nuevo currículo propuesto en el marco de la Ley Orgánica para la Mejora de la Calidad

Educativa (LOMCE), (MECD, 2015) incluye criterios de evaluación y estándares de aprendizaje evaluables ligados específicamente al tema, que serán referentes en la planificación de la concreción curricular y en la programación didáctica, los resultados que se obtienen de este trabajo pueden servir también para planificar la enseñanza según la nueva normativa. Además, puesto que las PAU sólo estarán vigentes algunos años más, nuestro estudio puede orientar en la elaboración de las pruebas de final de Bachillerato previstas en la LOMCE.

Marco Teórico

Nos apoyamos en algunos elementos del Enfoque Onto-Semiótico (GODINO; BATANERO; FONT, 2007), que asume que los objetos matemáticos (en nuestro caso los objetos relacionados con la inferencia estadística) surgen de las prácticas matemáticas realizadas para resolver problemas relacionados con el objeto. Ejemplos de problemas relacionados con la inferencia serían determinar el tamaño de muestra necesario para llevar a cabo una estimación por intervalos con una precisión dada o contrastar una hipótesis sobre la media de una población. Mediante la reiteración de estas prácticas en muchos problemas, se llega a configurar el significado del objeto matemático o conjunto de prácticas asociadas al mismo. De ello se deduce la importancia que cobran los problemas en el aprendizaje de un determinado tema. También se diferencia en este marco teórico entre el significado institucional (compartido dentro de una institución) y el significado personal (que adquiere una persona, por ejemplo, un estudiante y puede ser diferente al aceptado dentro de la institución).

Dentro de la institución de enseñanza, se diferencian: (1) el significado global, sería el más amplio, en nuestro caso, el significado de la inferencia en la propia estadística; (2) el significado pretendido del objeto dentro de un nivel educativo, sería el fijado por la institución, en nuestro caso, el marcado por las orientaciones curriculares; (3) el significado implementado, es el que se presenta a los estudiantes en cada centro o aula; y (4) el significado evaluado, que sería el contenido en las pruebas de evaluación.

El objetivo de nuestro trabajo es determinar el significado evaluado de la inferencia estadística en las Pruebas de Acceso a la Universidad de la asignatura de Matemáticas

Aplicadas a las Ciencias Sociales II y comparar con el significado pretendido en las orientaciones curriculares. La importancia de este punto se fundamenta en que las pruebas de evaluación nos dan acceso a la comprensión del tema por parte del estudiante. Como sugiere Godino (1996), la comprensión de un determinado objeto matemático por parte de los estudiantes no puede observarse directamente pues es un constructo psicológico inobservable. Sin embargo, puede ser evaluada indirectamente a través de las respuestas de los estudiantes a los ítems, tareas o pruebas de evaluación, es decir, las soluciones finales; estrategias; argumentos; símbolos usados; etc. Es fundamental entonces que las pruebas de evaluación sean válidas, es decir, que exista una correspondencia entre el significado institucional pretendido y el evaluado para un contenido. Esta correspondencia es la que tratamos de evaluar en la investigación para el contenido de inferencia.

Antecedentes

Al analizar los problemas de inferencia estadística en las pruebas PAU nos encontramos dos tipos de referentes para nuestro trabajo: a) Análisis de pruebas de evaluación; y b) Estudios sobre dificultades que pueden tener los estudiantes al resolver problemas de inferencia estadística.

a) Estudios de pruebas de evaluación

Entre los investigadores que se han interesado por analizar pruebas de evaluación encontramos a Rico (2006), quien relaciona el marco teórico de las pruebas de matemáticas PISA del año 2003 con la resolución de problemas desde el punto de vista curricular a través del concepto de competencia matemática, utilizando un modelo funcional sobre el aprendizaje de las matemáticas. Dicho modelo analiza el patrón a seguir a la hora de resolver una tarea matemática: campos de actuación (fenómenos, situaciones o contextos y problema que plantea la tarea); herramientas o matemáticas puestas en juego para resolverla (estructuras conceptuales y procedimientos) y procesos cognitivos que se movilizan para su ejecución (pensar y razonar, comunicar, justificar, representar, modelizar, plantear y resolver problemas).

Caraballo (2010) analiza 173 ítems incluidos en las pruebas de diagnóstico correspondientes al segundo curso de Educación Secundaria Obligatoria de cinco comunidades autónomas durante el curso académico 2008-2009. El objetivo principal de su investigación fue estudiar la correspondencia entre los instrumentos de evaluación utilizados por las comunidades autónomas para realizar estas pruebas y el modelo PISA. Para ello analizó en los problemas las siguientes variables: contenido matemático; el contexto o situación; el nivel de complejidad y los procesos que se deben utilizar para resolver una tarea matemática. La autora concluye la investigación señalando el desequilibrio existente en la distribución de los ítems respecto al contexto y al contenido. En cuanto al contexto, se muestra irrelevante o camuflado (ofrece apariencia del mundo real pero como pretexto ya que su relevancia no se mantiene durante el proceso de resolución) en más de una cuarta parte de los ítems analizados. En cuanto al contenido, en aproximadamente una tercera parte de los ítems se presentan tareas que en lugar de movilizar competencias, atienden a contenidos curriculares. Además, observa que habitualmente la información viene dada mediante representaciones gráficas o simbólicas.

Por otro lado, Castellanos (2013) realiza un estudio sobre las tablas y gráficos estadísticos de las pruebas SABER de Colombia para estudiantes de último grado del ciclo de educación básica primaria en los años 2003, 2006 y 2009. La autora analiza el tipo de gráfico o tabla utilizado (como por ejemplo el diagrama de sectores o el pictograma); la competencia evaluada (razonamiento, resolución o comunicación); nivel de lectura de gráficos (uso del gráfico para resolver la tarea); la actividad solicitada (interpretar una representación, organizar y representar datos) y los niveles de complejidad semiótica del gráfico (entre los descritos por Arteaga, 2011). En la misma línea, pero centrándose en las Pruebas de Diagnóstico Andaluzas Obligatorias para niños de 10 años, Mingorance (2014) analiza los gráficos y tablas estadísticas observando una alta frecuencia de los gráficos de barras frente a la escasez de todos los recomendados en el currículo.

Respecto a las pruebas PAU, aunque se han realizado diversas investigaciones, pocas han estado centradas en el análisis de los problemas planteados. Respecto a las pruebas de matemáticas, se han analizado los resultados de los estudiantes (ej., BLÁZQUEZ; LUENGO, 1989), la opinión de los profesores (RUIZ; DÁVILA; ETXEBERRÍA; SARASUA, 2013) o las características psicométricas de las prueba (ej., GAVIRIA, 2005). La más cercana a nuestro trabajo es la realizada por Espinel, Ramos y Ramos (2009) que analizan las dificultades encontradas por una muestra de estudiantes al resolver dos problemas de contraste de hipótesis propuestos en la comunidad canaria.

Por nuestra parte, en trabajos previos (e.j., LÓPEZ-MARTÍN, CONTRERAS, BATANERO; CARRETERO, 2015) hemos analizado los problemas de probabilidad propuestos en las Pruebas de Acceso a la Universidad en Andalucía en el periodo 2003-2014. Los resultados mostraron un fuerte énfasis en la probabilidad condicional, frente a la simple o compuesta, y en general, revelaban la dificultad de los problemas, con predominio de espacios muestrales compuestos, sucesos no equiprobables y datos dados en forma de probabilidad o porcentaje.

La investigación actual trata de complementar los anteriores para ofrecer una visión de la evaluación dedicada al Bloque de Estadística y Probabilidad dentro de las Pruebas PAU en Andalucía. Seguidamente comentamos algunas investigaciones relacionadas con las dificultades de los estudiantes en el tema de inferencia, que permiten valorar mejor la complejidad de los ejercicios analizados.

b) Dificultades de los estudiantes con la inferencia

Dentro de la investigación sobre didáctica de la estadística se han descrito bastantes errores e interpretaciones incorrectas relacionadas, en general, con la comprensión del contraste de hipótesis (FALK; GREENBAUM, 1995; BATANERO, 2000; HARRADINE; BATANERO; ROSSMAN, 2011). Estas dificultades se relacionan en parte con la comprensión de la probabilidad condicional, pues algunos conceptos de inferencia se definen usándola. Por ejemplo, el nivel de significación, denotado habitualmente como α, se define como la probabilidad de rechazar la hipótesis nula (H_0), supuesta cierta, es decir, $\alpha = P(\text{rechazar } H_0 | H_0 \text{ es cierta})$. Por ello, si teniendo que efectuar un gran número de contrastes de hipótesis se trabaja a un nivel de significación del 5%, esto quiere decir que si H_0 es cierta, la rechazamos 5 de cada 100 veces. Muchos estudiantes intercambian la condición y el suceso en esta probabilidad condicional, interpretando como la probabilidad de que la hipótesis nula sea cierta, una vez que la decisión de rechazarla se ha tomado, esto es, suponer que $\alpha = P(H_0 \text{ es cierta} | \text{ se ha rechazado } H_0)$. Así, Vallecillos (1999) encontró este tipo de error en una investigación con más de 400 estudiantes en la Universidad de Granada, repitiéndose en estudiantes de diferentes especialidades (por ejemplo, de medicina, ingeniería o psicología).

Otro error común en los contrastes de hipótesis es el planteamiento de la hipótesis nula y alternativa. Es decir, algunos estudiantes piensan que la hipótesis nula es la que queremos demostrar (no la que queremos rechazar). Posiblemente se deba esta creencia a que en matemáticas casi siempre se trata de probar una hipótesis (aunque en el método de reducción al absurdo se trata de rechazarla). Vallecillos (1999) describió dos creencias erróneas sobre el tipo de prueba que proporciona el contraste de hipótesis:

- *El contraste de hipótesis es una prueba probabilística de la hipótesis* y por tanto permite calcular la probabilidad de que una hipótesis sea cierta. Según Batanero (2000), este cálculo solo es posible en inferencia bayesiana, pero no tiene sentido en inferencia frecuencial.
- *El contraste estadístico es un método matemático*; como tal, y al ser la matemática una ciencia exacta, al finalizar hemos probado la verdad o falsedad de una hipótesis. Esta creencia es siempre errónea; la tienen algunos estudiantes que tienen poca base matemática y una mirada determinista de su entorno. Suelen tener dificultades de comprensión, aprenden el cálculo de memoria y piensan que el resultado debe ser cierto o falso. Falk y Greenbaum (1995) sugieren que esta creencia se debe a la existencia de mecanismos psicológicos; algunas personas desean minimizar la incertidumbre en la decisión y suponen que la obtención de un resultado estadísticamente significativo les ayuda a ello.

La creencia de que rechazar la hipótesis nula supone demostrar que es errónea también se encontró en la investigación de Saldanha y Thompson (2002), quienes indican que las ideas de probabilidad y atipicidad son fundamentales para comprender la lógica de la prueba de hipótesis, donde se rechaza una hipótesis nula cuando una muestra de esta población se considera lo suficientemente atípica si la hipótesis nula es cierta. También hay que comprender el muestreo como un sistema de ideas interrelacionadas, selección al azar, replicación, variabilidad y distribución. Mientras que comprender la idea de muestra aleatoria simple es fundamental, probablemente es más importante entender que cada muestra es sólo una de las posibles entre las que podrían haberse elegido. Otros errores descritos por Espinel, Ramo y Ramos (2009) en su estudio es permutar las hipótesis nula y alternativa o confundir el valor estimado de un parámetro con el hipotético, cálculo erróneo de probabilidades y confusión de nivel de significación y punto crítico.

Aunque los errores descritos aparecen principalmente en el contraste de hipótesis, también es posible encontrar, pero no con tanta frecuencia, errores que se producen al interpretar una probabilidad condicional en el cálculo o interpretación de intervalos de confianza (CUMMING; WILLIAMS; FIDLER, 2004). Hay que añadir también las posibles dificultades en diferenciar la distribución estadística (datos), la distribución de la variable aleatoria (población) y la distribución muestral, que se deben manejar simultáneamente en el trabajo con inferencia (SCHUYTEN, 1991).

En definitiva, la resolución de un problema de inferencia plantea un reto importante a los estudiantes. Es por ello importante determinar claramente cuál es el contenido de los problemas sobre inferencia en las Pruebas de Acceso a la Universidad, entre otras razones, por la influencia que la puntuación obtenida en las mismas tiene sobre las posibilidades del estudiante para acceder a una titulación de su elección. En lo que sigue se describe la metodología y resultados de nuestro estudio.

Metodología

Las Pruebas de Acceso a la Universidad (PAU)

Las actuales Pruebas de Acceso a la Universidad están regidas por el Real Decreto 1892/2008, por el que se regula las condiciones para el acceso a las enseñanzas universitarias oficiales de grado y los procedimientos de admisión a las universidades públicas españolas (MP, 2008). A su vez, este decreto procede de la Ley Orgánica de Educación (LOE), que exige, en su artículo 38, la superación de una prueba de madurez que permita valorar los conocimientos adquiridos en Bachillerato y la capacidad de los estudiantes para iniciar estudios universitarios.

La prueba correspondiente a Matemáticas Aplicadas a las Ciencias Sociales II evalúa los conocimientos y capacidades de los estudiantes según la que lleva el mismo nombre, cuyos contenidos se indicaron en la introducción de este trabajo. En Andalucía esta prueba consta de dos opciones (opción A y opción B), cada una de ellas formada por cuatro ejercicios y donde el estudiante, bajo su parecer, elige realizar íntegramente una de ellas. Tanto la opción A como la opción B están divididas en tres tipologías de ejercicios. El primer ejercicio pertenece al bloque de contenidos de Números y Álgebra, el segundo al bloque de Análisis, y los ejercicios tercero y cuarto corresponden al Bloque de Estadística y Probabilidad, más concretamente, a Probabilidad e Inferencia Estadística. En la actualidad, los ejercicios incluidos en cada opción tienen asignada una puntuación máxima de 2,5, luego, el 50% de la nota total de la prueba corresponde al Bloque de Estadística y Probabilidad. En lo que se refiere a nuestro análisis, dentro del Bloque de Estadística y Probabilidad, nos restringimos a uno de estos ejercicios, que corresponde a los contenidos relacionados con la inferencia estadística.

La investigación llevada a cabo es esencialmente de tipo cualitativa, ya que se basa en el análisis de contenido en un documento (la colección de problemas propuestos en las pruebas de selectividad). Esta técnica supone que un texto puede dividirse en unidades que pueden clasificarse en un número reducido de categorías, en función de variables subyacentes, y que permiten realizar inferencias sobre su contenido (KRIPPENDORFF, 1997). En nuestro caso se ha analizado el contenido de todos los problemas, presentando la resolución de algunos problemas tipo y a partir de ellos se deduce el contenido requerido en su resolución. Este análisis se complementa con un estudio estadístico de los problemas, de tipo cuantitativo.

Según Bisquerra (1989), el proceso de investigación que se ha desarrollado es inductivo, pues se parte de casos particulares (ejemplos de pruebas PAU) con el fin de obtener generalizaciones a partir de estas observaciones. La investigación realizada es aplicada, descriptiva y exploratoria, ya que está encaminada a obtener criterios de mejora para la evaluación de los estudiantes, sin la realización de manipulaciones sobre las variables y sin considerar hipótesis de partida.

Muestra de problemas analizados

La muestra considerada en esta investigación está formada por todos los problemas propuestos en las pruebas PAU de la modalidad de Matemáticas Aplicadas a las Ciencias Sociales II en la Comunidad Autónoma de Andalucía desde 2003 a 2014, en total 12 años. En cada uno de estos años se revisaron las seis pruebas disponibles (tres tipos de examen diferentes, con dos opciones a elegir para cada uno y en dos convocatorias diferentes: julio y septiembre) cada una las cuales contiene un ejercicio de inferencia estadística (12 problemas por año). Así pues, el tamaño total de la muestra analizada es de 144 problemas. Mediante un análisis semiótico, según el marco teórico descrito anteriormente, (GODINO; BATANERO; FONT, 2007), se ha analizado el contenido matemático, determinando los objetos matemáticos implícitos en la resolución de cada problema.

Es importante señalar que con esta investigación no se pretende extrapolar los resultados a otras pruebas diferentes a las analizadas, ya que la muestra con la que se está trabajando es de tipo intencional, algo propio de la metodología cualitativa. No obstante, pensamos que las conclusiones pueden servir para conjeturar hipótesis provisionales sobre el contenido de inferencia en las pruebas realizadas en otros años o en otras comunidades autónomas, aportando resultados que serían necesarios para poder contrastar dichas hipótesis.

Variable analizada

Nuestro estudio comenzó clasificando los contenidos recogidos en el Decreto de Educación Secundaria y Bachillerato (MECD, 2015), que fueron presentados en la introducción del presente trabajo. También se resolvieron los problemas propuestos para analizar su contenido matemático. A partir de estos análisis se han determinado los diferentes campos de problemas específicos que se plantean y que define el contenido matemático de la tarea. El campo de problemas es un elemento fundamental en el Enfoque Onto-Semiótico (GODINO; BATANERO; FONT, 2007), pues determina las prácticas y objetos matemáticos involucrados en la actividad matemática y por tanto, en el proceso de evaluación de las pruebas PAU. También se considera como parte del análisis los campos de actuación considerados por Rico (2006). Más concretamente, se han diferenciado tres campos principales de problemas: A) Muestreo; B) Intervalos de confianza, y C) Contraste de hipótesis.

Los dos primeros se han subdividido a su vez en dos subcampos de problemas, que serían A1) Enumerar las diferentes muestras de una población o realizar cálculos estadísticos de las distribuciones muestrales en poblaciones finitas; A2) Identificar la distribución muestral de un estadístico o realizar cálculo de probabilidades con la misma; B1) Calcular o interpretar un intervalo de confianza; y B2) Relación entre confianza, error y tamaño muestral. A continuación se presentan los resultados, describiendo con más detalle las categorías a partir de la solución de un ejemplo. Seguidamente presentamos una síntesis del contenido matemático de las tareas y correspondencia con los estándares de evaluación evaluables fijados en el nuevo currículo (MECD, 2015) para la inferencia. Además, se presenta la distribución, globalmente y por año, de las categorías en la muestra de problemas.

Resultados

Campos de problemas propuestos y contenido evaluado en cada uno de ellos

Se han analizado los distintos tipos de problemas sobre inferencia que se proponen en cada una de las pruebas PAU de la muestra. Los contenidos se han dividido en tres campos principales: Muestreo (con los apartados A1 y A2), Intervalos de confianza (divididos en las categorías B1 y B2) y Contraste de hipótesis (categoría C). A continuación se describe con detalle esta clasificación, analizando un ejemplo en cada categoría. En cada una de ellas se puede observar la diversidad y especificidad de objetos matemáticos que las determinan y por tanto, tienen diferente grado de dificultad para el estudiante.

A1 Composición de las muestras y cálculos estadísticos de las distribuciones muestrales en poblaciones finitas

Dentro de este bloque se han considerado aquellos ítems relacionados con la composición o enumeración de las muestras de una población finita en unas condiciones dadas o cálculos de sus estadísticos. Los datos que proporciona el problema, generalmente, son el tipo de población, el tamaño de la muestra y las condiciones del muestreo aleatorio (simple, sistemático y estratificado). Un ejercicio de esta categoría se analiza a continuación; en el primer apartado, partiendo de una población finita discreta, se extraen aleatoriamente todas las muestras de un tamaño determinado, y se pide calcular para cada una de ellas el estadístico de interés. En el segundo apartado se pide la composición de una muestra de tamaño 40 mediante un muestreo aleatorio estratificado con afijación proporcional.

Ítem 3A (2014)

a. Determine todas las muestras de tamaño 2 que, mediante muestreo aleatorio simple, se pueden extraer del conjunto {6, 9, 12} y calcule la varianza de las medias de estas muestras.

b. Una empresa fabrica cuatro productos A, B, C y D, de los que elabora diariamente 40, 15, 25 y 120 unidades respectivamente. Si un día se quiere elaborar una muestra de 40 unidades con los productos fabricados, por muestreo aleatorio estratificado con afijación proporcional, ¿qué número de unidades de cada producto se debe elegir?

En relación a la resolución del primer apartado es necesario comprender el tipo de muestreo que se está realizando, ya que el número de muestras obtenidas mediante un muestreo aleatorio con o sin reemplazamiento no coinciden. Dado que, en el enunciado del ejercicio, no se especifica lo contrario, se supone que el muestreo es con reemplazamiento, es decir, es posible extraer una muestra con elementos repetidos. Una vez aceptado este hecho se empezará a enumerar las posibles muestras de 2 elementos, obteniendo como resultado en total 9 muestras: (6, 6), (6, 9), (6, 12), (9, 6), (9, 9), (9, 12), (12, 6), (12, 9) y (12, 12).

Para obtener la varianza de las medias muestrales, se requiere implícitamente el concepto de distribución muestral de la media, que se debe formar obteniendo para cada una de las muestras sus medias. A partir de dicha distribución muestral, la varianza se obtiene como $\sigma_{\bar{x}}{}^2 = \frac{\sum (x_i - \mu_{\bar{x}})^2 n_i}{n}$ o equivalentemente $\sigma_{\bar{x}}{}^2 = \frac{\sum \bar{x}_i{}^2 n_i}{n} - (\mu_{\bar{x}})^2$ siendo $\bar{x}_i$ los valores de las medias en cada muestra y n_i la frecuencia absoluta de los mismos. Para facilitar el cálculo, además de la calculadora, se puede recurrir al uso de una tabla estadística, véase Tabla 1. A partir de ella, se tiene que la media y la varianza de la distribución muestral son respectivamente $\mu_{\bar{x}} = \frac{\sum \bar{x}_i n_i}{n} = \frac{81}{9} = 9$ y $\sigma_{\bar{x}}{}^2 = \frac{756}{9} - (9)^2 = 3$.

Tabla 1: Organización de cálculos en la resolución del problema 3A (2014)

Medias muestrales ($\bar{x}_i$)	$\boldsymbol{n_i}$	$\boldsymbol{\bar{x}_i n_i}$	$\boldsymbol{\bar{x}_i{}^2 n_i}$
6	1	6	36
7,5	2	15	112,5
9	3	27	243
10,5	2	21	220,5
12	1	12	144
Sumas	9	81	756

Fuente: Elaboración Propia

Otra forma de actuar es basándose en el Teorema Central del Límite. Sabiendo la relación existente entre la media y desviación de una población, y la distribución de las medias de las muestras de tamaño *n* obtenidas mediante muestreo aleatorio simple con reemplazamiento, se obtendría que la distribución de las medias de las muestras es igual a la media de la población ($\mu=9$) y su varianza es igual al cociente entre la varianza de la población y el tamaño de las muestras obtenidas:

$$\sigma_{\bar{x}}^{2}=\frac{\sigma^{2}}{n}=\frac{\frac{6^{2}+9^{2}+12^{2}}{3}-9^{2}}{2}=3$$

En la resolución del apartado b) es necesario distinguir los diferentes tipos de muestreos aleatorios que se han estudiado en los contenidos de la asignatura (simple, estratificado, sistemático). Al tratarse de un muestreo aleatorio estratificado con afijación proporcional, existe una relación de proporcionalidad entre el número total de elementos de todos los estratos ($N=N_1+N_2+\ldots+N_k$, donde N_i es el tamaño del estrato i) y el número total de elementos de las muestra extraída ($n=n_1+n_2+\ldots+n_k$, siendo n_i el tamaño de la muestra i correspondiente al estrato i) es decir, $\frac{n_1}{N_1}=\frac{n_2}{N_2}=\cdots=\frac{n_k}{N_k}=\frac{n}{N}$.

En este caso, ha de verificarse que $\frac{n_1}{40}=\frac{n_2}{15}=\frac{n_3}{25}=\frac{n_4}{120}=\frac{40}{40+15+25+120}=\frac{40}{200}$. A partir de la expresión dada se tiene que $\frac{n_1}{40}=\frac{40}{200}$ y por tanto $n_1=8$. Procediendo de forma análoga se tiene que $n_2=3$, $n_3=5$ y $n_4=24$.

Los contenidos implicados en la tarea analizada son esenciales para el razonamiento posterior sobre las características de una población (tamaño excesivamente elevado provocando una imposibilidad material y temporal de muestrear todos sus individuos). En la práctica se recurre a una muestra para inferir datos de esa población. Por consiguiente, es necesario garantizar la representatividad de la muestra prestando atención al tamaño muestral y al proceso de selección de los elementos que la forman. Por tal motivo, la selección de la muestra juega un papel sumamente importante a la hora de realizar inferencia sobre los datos descriptivos de una población.

A2 Distribución muestral en poblaciones infinitas y cálculo de probabilidades

En la clasificación realizada se han considerado, dentro de esta categoría, todos aquellos problemas en los que se pide describir la distribución muestral de un cierto estadístico en poblaciones infinitas, lo que requiere que el estudiante utilice algún modelo teórico de distribución. Como dato del problema se da el tipo de población y el estadístico de interés. Entre los estudiados a nivel de Bachillerato, los estadísticos de uso son la media muestral y la proporción muestral, con distribuciones normal y binomial respectivamente, aproximándose ambas a la normal en muestras grandes.

También se da el tamaño de la muestra, y en caso de analizar la media, la desviación típica de la población, mientras que si se trata de la proporción, el estudiante ha de recordar su fórmula. Se supone muestreo aleatorio con reemplazamiento, lo que implica que es posible obtener, como resultado de un número infinito de muestras aleatorias independientes de tamaño *n* proveniente de la población de interés, la distribución del estadístico. Una vez determinada la distribución del estadístico en estudio, se suele pedir calcular probabilidades, como en el ejemplo que se analiza a continuación.

Ítem 5B (2011)
Sea *X* una variable aleatoria Normal de media 50 y desviación típica 4. Se toman muestras de tamaño 16.
a. ¿Cuál es la distribución de la media muestral?
b. ¿Cuál es la probabilidad de que la media muestral esté comprendida entre 47,5 y 52,5?

Basándose en el Teorema Central del Límite se obtiene la resolución del primer apartado. Recordemos que dicho teorema establece que si la población en estudio tiene media μ, desviación típica σ y el tamaño de la muestra (n) es superior a 30, entonces la distribución de la media muestral: $\bar{X}$ es normal con igual media y desviación típica el cociente entre la desviación típica de la población y la raíz cuadrada del tamaño de la muestra.

En nuestro caso: $\bar{X} \to N\left(50, \frac{4}{\sqrt{16}}\right) = N(50,1)$.

El hecho de describir la distribución de la media de las muestras nos permite obtener probabilidades sobre dicha variable, aplicando los procedimientos de tipificación y lectura de tablas de probabilidades que se conocen. Así pues, tipificando la distribución de la media muestral $\left(Z = \left(\bar{X} - 50\right)/1\right)$, se obtiene la distribución normal tipificada y el estudiante puede consultar las tablas de probabilidades para la resolución del problema.

$$P\left\{47,5\leq \bar{X}\leq 52,5\right\}=P\left\{\frac{47,5-50}{1}\leq\frac{\bar{X}-50}{1}\leq\frac{52,5-50}{1}\right\}=P\left\{-2,5\leq Z\leq 2,5\right\}$$
$$=P\left\{Z\leq 2,5\right\}-P\left\{Z\leq -2,5\right\}=P\left\{Z\leq 2,5\right\}-\left(1-P\left\{Z\leq 2,5\right\}\right)=0,9876.$$

Por tanto, la probabilidad de que la media muestral esté comprendida entre 47,5 y 52,5 es 0,9876, es decir, en el 98,76 % de las muestras de tamaño 16 la media estará en este intervalo.

Como en el campo de problemas A1, en los problemas clasificados como A2 el estudiante debe tener en cuenta la relación existente entre la media y varianza de la población y las muestras de tamaño determinado e identificar correctamente la distribución de la variable en estudio. Además, debe ser capaz de tipificar y leer correctamente las tablas de probabilidades de una normal tipificada con el fin de obtener la probabilidad pedida. Igualmente ha de aplicar reglas básicas de cálculo de probabilidades en un intervalo según una variable aleatoria continua.

B1 Cálculo o interpretación de un intervalo de confianza

Dentro de este bloque se han considerado tanto los problemas en los que se pide explícitamente la construcción de un intervalo de confianza para un cierto estadístico (generalmente la media o la proporción muestral) así como los relacionados con la interpretación de la información aportada por el intervalo. En lo que sigue analizamos un ejemplo.

Ítem 1A (2006)

De 500 encuestados en una población, 350 se mostraron favorables a la retransmisión de debates televisivos en tiempos de elecciones. Calcule un intervalo de confianza, al 99,5 %, para la proporción de personas favorables a estas retransmisiones.

En la resolución del ejercicio el estudiante debe tener en cuenta que para estimar el parámetro de una población mediante intervalo de confianza, en nuestro caso la proporción poblacional, debe considerarse como estimador puntual la proporción muestral, así como su distribución.

Cuando se trata de determinar por intervalo de confianza (o realizar un contraste de hipótesis) la proporción de individuos que posee una característica o atributo, se utiliza la distribución binomial, y esta puede ser aproximada por una distribución normal en ciertos casos. Concretamente, si en una población la proporción de individuos que posee un atributo es p, entonces, la proporción muestral de individuos con esta característica en las muestras de tamaño n es $\hat{p}$, y siempre que el tamaño de la muestra sea superior a 30, se distribuye aproximadamente según una distribución normal de media p y desviación típica $\sqrt{\frac{p(1-p)}{n}}$.

Como consecuencia, al tipificar: $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \rightarrow N(0;1)$.

Un elemento básico en la construcción del intervalo es el coeficiente de confianza $(1-\alpha)100\%$. Dado este coeficiente, se desean encontrar los límites de un intervalo de valores $[e_1, e_2]$, tal que $P(e_1 \leq p \leq e_2) = 1-\alpha$. Para calcularlo, se comienza identificando un intervalo de valores de la distribución normal tipificada, centrado en el origen y tal que el área de cada cola sea $\frac{\alpha}{2}$, es decir, $P\{-k < Z < k\} = 1-\alpha$. Realizando una lectura inversa de la tabla de la distribución se tiene que, los valores de $-k$ y k son aquellos que verifican que $P\{Z \leq -k\} = \alpha/2$ y $P\{Z \leq k\} = 1-\alpha/2$, respectivamente. A partir de estos valores se despejan los extremos del intervalo de confianza de la variable Z.

En nuestro caso, usando la tipificación de la proporción muestral, se obtiene: $P\left\{|\hat{p}-p| < k\sqrt{\frac{p(1-p)}{n}}\right\} = 1-\alpha$, de donde se despejan los extremos del intervalo de confianza. Puesto que estos extremos dependen del parámetro poblacional que es desconocido, una vez obtenida la proporción de individuos que poseen una cualidad, se empleará dicho valor muestral como el estimador de p.

Así pues, dado que $k = Z_{1-\frac{\alpha}{2}}$, se concluye que el intervalo de confianza de la proporción poblacional con un nivel de confianza $(1-\alpha)100\%$ es $\left(\hat{p} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$.

revista
PARANAENSE
DE EDUCAÇÃO
MATEMÁTICA

Puesto que $k = Z_{0,9975} = 2,81$ y $\hat{p} = \frac{350}{500} = 0,7$ obtenemos que el intervalo de confianza solicitado es I.C.$_p$ = (0,6421, 0,7576). Hay que destacar que una vez que se ha calculado el intervalo para una muestra determinada, no es correcto interpretar que la probabilidad de que el parámetro pertenezca al intervalo es $(1-\alpha)$ ya que, una vez obtenido deja de ser aleatorio y la probabilidad será 1 si el intervalo es de los $(1-\alpha)100\%$ que contienen al parámetro, o 0 si el intervalo es uno de los α100% intervalos que no contienen al parámetro. Por tanto, no tiene sentido hablar de probabilidad sino de confianza. La confianza está puesta en que el método de construcción de los intervalos nos asegura que (1 − α)100% de las muestras producirán intervalos que contienen al parámetro.

B2 Relación entre confianza, error de estimación y tamaño muestral

En este grupo se han considerado todos los ejercicios en los que se solicita calcular el coeficiente de confianza, error de estimación o tamaño muestral después de haber fijado dos de ellas. Como datos se da la población y estadístico considerado o bien el intervalo de confianza. Un ejemplo se muestra a continuación.

Ítem 3B (2007)
Se sabe que (45,13, 51,03) es un intervalo de confianza, al 95%, para la media de una variable aleatoria que sigue una distribución Normal con desviación típica 15.
a. ¿Cuál es el error cometido?
b. Calcule, con el mismo nivel de confianza, el tamaño muestral mínimo necesario para que el error no sea superior a 1,8.

En este problema el estudiante debe, en primer lugar, diferenciar entre estadístico (media de la muestra) y parámetro (media de la población). Debe asimismo comprender el significado del intervalo de confianza (un rango de posibles valores para el parámetro, en este caso la media de la población) y del coeficiente de confianza (95% de los intervalos construidos de la misma población con muestras del mismo tamaño contendrán al parámetro; pero no sabemos si el intervalo particular lo contiene o no). Puesto que se da la distribución de la población como dato, y se trata de una distribución normal $N(\mu,\sigma)$ entonces se verifica que la variable aleatoria media muestral se distribuye según una distribución normal:

$N\left(\mu,\frac{\sigma}{\sqrt{n}}\right)$, donde *n* es el tamaño de la muestra, que no se da en este problema.

Para resolver el primer apartado, el estudiante ha de recordar que el error de estimación se define como la mitad de la amplitud del intervalo, es decir, es la mitad de la diferencia existente entre el extremo superior e inferior del intervalo; por tanto, $E=\frac{1}{2}\times(51{,}03-45{,}13)=2{,}95$.

Si nos centramos en la resolución del segundo apartado, al pretender disminuir el error (manteniendo el nivel de confianza) la única variable que puede ser modificada es el tamaño muestral, ya que el error de estimación para el caso de la distribución normal, que es la que utiliza en estos problemas, es equivalente a $E=Z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$. Dada la relación existente entre el error de estimación y el tamaño muestral, se puede establecer que cuanto mayor sea el tamaño de la muestra la precisión en la estimación realizada será mejor, ya que el intervalo es más estrecho. En nuestro caso, dado que el nivel de confianza es del 95% (entonces $Z_{1-\alpha/2}=1{,}96$), la desviación típica de la población es 15 y el error de estimación debe ser no superior a 1,8 entonces concluimos que al menos la muestra debe estar formada por 267 elementos.

A partir de la expresión del error, el estudiante puede razonar a través de las siguientes propiedades: a) a medida que aumenta el tamaño de la muestra (manteniendo constante el nivel de confianza) el error de estimación disminuirá; y b) manteniendo constante el tamaño de la muestra, el error aumentará a medida que aumente el nivel de significación. Observemos que una forma más coherente para conseguir mayor precisión en la estimación es el aumento del tamaño de la muestra, ya que habitualmente no es recomendable utilizar un nivel de confianza inferior al 90%.

C Test de hipótesis

En el último tipo de problemas se pide al estudiante comprobar si una afirmación estadística relativa al valor de un parámetro de la población es compatible con el resultado obtenido de la información aportada por una muestra aleatoria simple procedente de la misma. Las diferencias que pueden presentarse es el tipo de población considerada (normal o binomial) y los apartados del problema o los datos aportados. Tomemos como ejemplo el siguiente, propuesto en Junio de 2013 (prueba de reserva a):

Ítem 5A (2013)
Un director sanitario sostiene que el Índice de Masa Corporal (IMC) medio de los adolescentes de su distrito no supera el nivel 25 (sobrepeso). Para contrastar su afirmación toma una muestra aleatoria de 225 adolescentes que da como resultado un IMC medio de 26. Sabiendo que el IMC sigue una distribución Normal con desviación típica 5 discuta, mediante un contraste de hipótesis con $H_0 \equiv \mu \leq 25$, si la afirmación del director sanitario es correcta, con un nivel de significación del 5%.

El problema pide contrastar una hipótesis sobre la media de una población normal con desviación típica conocida $N(\mu,5)$. Se trata de un contraste unilateral que se da en el enunciado $H_0 : \mu \leq 25$. Se pide tomar una decisión (rechazar o no la hipótesis con un nivel de significación dado); para ello, se debe determinar las regiones de aceptación y rechazo en el contraste. Para la resolución del problema, el estudiante debe recordar los siguientes conceptos y propiedades:

- En la realización de un contraste de hipótesis sobre la media de la población (μ), ha de utilizar un estimador, en este caso la media muestral. El valor de la misma se debe identificar de los datos del problema.
- Si la variable aleatoria sigue una distribución normal $N(\mu,\sigma)$, la distribución de las medias de las muestras de tamaño 225 es $N\left(\mu,\frac{\sigma}{\sqrt{n}}\right)=N\left(\mu,\frac{5}{\sqrt{225}}\right)=N\left(\mu,\ 0{,}33\right)$.
- Suponiendo cierta la hipótesis nula, es decir, $\mu = 25$, el estadístico de contraste se obtiene de tipificar la variable media muestral, es decir, $Z=\frac{\overline{X}-25}{1/3}$. Por tanto, considerando que la media de la muestra dada es 26 el valor dicho estadístico es 3.

En dicha distribución normal tipificada, el nivel de significación o equivalentemente la probabilidad de obtener un valor superior o igual a 3, suponiendo la hipótesis nula H_0 cierta, sería $P(Z \geq 3 \mid H_0) = 0{,}0013$. Así pues, si la hipótesis nula fuera cierta, es decir, si el índice medio de masa corporal en su distrito fuese igual o menor a 25, la probabilidad de obtener una muestra de 225 adolescentes con un índice medio igual o mayor a 26 sería 0,0013. En consecuencia, dado que 0,0013<0,05, lo razonable sería rechazar la afirmación del director sanitario, ya que los datos empíricos no apoyan la hipótesis nula planteada.

Síntesis del contenido evaluado en los diferentes campos de problemas

Para sintetizar, el Cuadro 1 contiene los resultados del análisis de la solución de un ejemplo de cada categoría de los diferentes campos de problemas. Observamos el dominio que se requiere de conceptos y procedimientos en cada uno de dichos problemas.

Cuadro 1: Contenidos matemáticos evaluados en los diferentes campos de problemas

	Contenidos	**A1**	**A2**	**B1**	**B2**	**C**
Conceptos y propiedades	Población y muestra	x	x	x	x	x
	Muestreo con y sin reemplazamiento: diferencias	x				
	Distribución muestral	x	x	x	x	x
	Relación entre estadísticos en la población y en la distribución muestral	x	x	x	x	x
	Modelo teórico de distribución (binomial o normal)		x	x	x	x
	Distribución normal tipificada		x	x	x	x
	Estimador de un parámetro			x	x	x
	Intervalo de confianza			x	x	
	Coeficiente de confianza; significado			x	x	
	Relación entre amplitud de intervalo, precisión y tamaño muestral				x	
	Hipótesis (nula y alternativa; diferencias)					x
	Estadístico de contraste					x
	Nivel de significación					x
	Regiones de aceptación y rechazo					x
	Lógica del contraste					x
	Probabilidad condicional	x		x	x	x
Procedimientos	Enumerar las posibles muestras de una población finita	x				
	Calcular estadísticos en las muestras	x				
	Calcular estadísticos de la distribución muestral	x	x			x
	Calcular tamaño de la muestra usando proporciones	x				
	Tipificación		x	x		
	Lectura de tablas de la distribución normal		x	x	x	x
	Cálculo de probabilidades en intervalos		x	x		x
	Cálculo de percentiles en la distribución normal			x		x
	Cálculo de extremos en intervalos de confianza			x		
	Cálculo de regiones de aceptación y rechazo					x
	Toma de decisión sobre una hipótesis					x

Fuente: Elaboración propia

Notamos conocimientos comunes evaluados en los diferentes campos de problemas considerados, en particular, los de población y muestra, distribución muestral y relación entre los estadísticos de la muestra, la población y la distribución muestral. Ello fuerza al estudiante a diferenciar los tres planos de distribución señalados por Schuyten (1991): la distribución de datos en la muestra, la distribución de la variable de interés en la población y la distribución muestral del estadístico en todas las posibles muestras obtenidas de la población. Esta diferencia (y de los correspondientes resúmenes como la media o la varianza) es compleja y puede llevar al estudiante a cometer errores; por ejemplo, a confundir la media poblacional hipotética y su estimador, como ocurrió en la investigación de Espinel, Ramos y Ramos (2009).

Todos los campos de problemas propuestos se apoyan fuertemente en la comprensión de la probabilidad condicional, que interviene tanto en la definición de las distribuciones muestrales (condicionadas al valor del parámetro), intervalos de confianza (condicionados con el valor del estadístico) como en el contraste de hipótesis (las regiones de aceptación y rechazo se calculan bajo la condición de ser cierta la hipótesis nula). Puesto que el concepto de probabilidad condicional es difícil y se han descrito números sesgos sobre el mismo (Ver DÍAZ; CONTRERAS; BATANERO; ROA, 2012) los profesores deberán prestar especial atención a su uso en inferencia, para asegurar la correcta solución de los estudiantes a los problemas propuestos en las pruebas de acceso.

Por otro lado, encontramos contenidos específicos en cada campo de problema definido, destacando los asociados, por un lado, a los campos B1 y B2 sobre intervalos de confianza y a C sobre contraste de hipótesis. En general, estos tres campos tienen mayor cantidad de contenidos evaluados en comparación con los de A1 y A2 sobre comprensión del muestreo y distribución muestral, a pesar de que la construcción de intervalos de confianza y el contraste de hipótesis se apoyan en este concepto. Esto implica una mayor complejidad de este tipo de problemas respecto a los referidos exclusivamente a muestreo y distribución muestral.

Comparación del significado pretendido y evaluado en los problemas

Como se indicó en la introducción, es importante asegurar la correspondencia entre el significado que el currículo propone para los objetos matemáticos (en este caso la inferencia) y el significado evaluado. Para estudiar esta correspondencia, en el Cuadro 2, realizamos una descomposición de los contenidos sobre inferencia propuestos para las Matemáticas Aplicadas a las Ciencias Sociales II en el Decreto de Enseñanza Mínimas (MEC, 2007, p. 45476).

En dicho cuadro observamos que todos los contenidos fijados en el significado previsto de la inferencia se evalúan en algunos de los campos de problemas planteados, aunque no con la misma intensidad. Así los contenidos relacionados con el Teorema Central del Límite y la aproximación de la distribución binomial a la normal se utilizará siempre que se trabaje con la distribución binomial para obtener distribuciones muestrales, calcular o interpretar intervalos de confianza y plantear contrastes de hipótesis. Los problemas relacionados con la elección de las muestra sólo se evalúan en los campos de problemas relacionados con ellas; el conocimiento sobre intervalo de confianza o contrate de hipótesis se evalúa en los campos de problemas B o C respectivamente y el conocimiento de la distribución muestral en todos los campos propuestos.

Cuadro 2: Descomposición del criterio de evaluación sobre inferencia para la asignatura Matemáticas Aplicadas a las Ciencias Sociales II (MEC, 2007, p.45477)

Descomposición de los criterios de evaluación	**A1**	**A2**	**B1**	**B2**	**C**
Implicaciones prácticas de los teoremas: Central del límite, de aproximación de la Binomial a la Normal y Ley de los Grandes Números.		x	x	x	x
Problemas relacionados con la elección de las muestras. Condiciones de representatividad.	x	x			
Parámetros de una población			x	x	x
Distribuciones de probabilidad de las medias y proporciones muestrales.	x	x	x	x	x
Intervalo de confianza para el parámetro p de una distribución binomial y para la media de una distribución normal de desviación típica conocida.			x	x	
Contraste de hipótesis para la proporción de una distribución binomial y para la media o diferencias de medias de distribuciones normales con desviación típica conocida					x

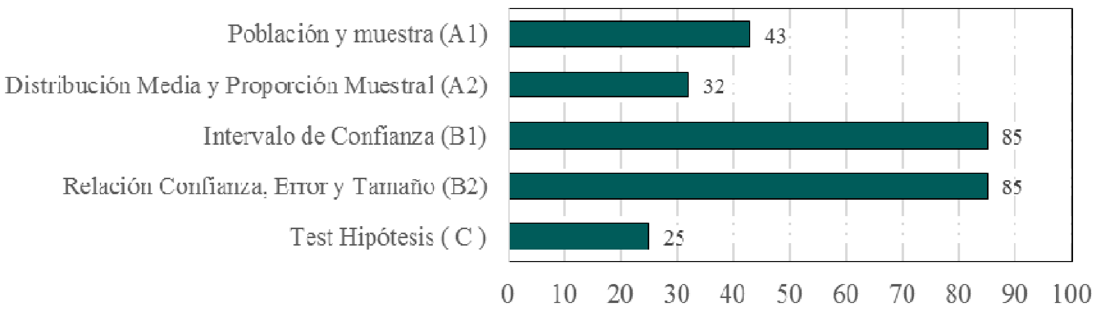
Fuente: Elaboración propia

Estudio estadístico de los problemas

Una vez llevado a cabo el análisis cualitativo, se realizó un estudio estadístico de los problemas para analizar la frecuencia con que se proponen lo diferente campos (y por tanto se evalúan los diferentes contenidos). En la Figura 1 se muestra la distribución global de los campos de problemas propuestos en los diferentes apartados de los 144 ejercicios analizados. Puesto que en el mismo ejercicio se pueden plantear preguntas de diferentes campos de problemas, el total supera bastante al número de ejercicios (271); prácticamente en todos los ejercicios se ha tratado de evaluar el conocimiento de más de un campo de problemas de entre los que hemos definido.

Se observa que los ejercicios que aparecen con mayor frecuencia en las pruebas son los relacionados con la construcción e interpretación de los intervalos de confianza y aquellos que hacen referencia a la relación entre la confianza, el error de estimación y el tamaño muestral. Conjuntamente, los ejercicios relacionados con el intervalo de confianza (campos B1 y B2) han sido un total de 170, por lo que se deduce la gran importancia dada a los intervalos en las PAU. En este sentido, y teniendo en cuenta que en estos problemas también se evalúa la comprensión de la distribución muestral y sus estadísticos, habría una carencia de problemas que evalúen la comprensión del contraste de hipótesis, en relación con los contenidos pretendidos (MEC, 2007).

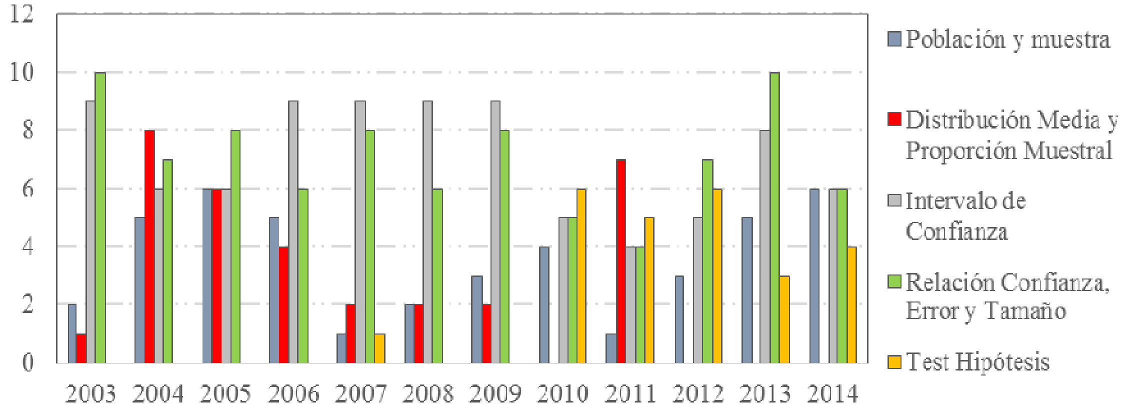
Figura 1: Frecuencia de ejercicios según campo de problemas



Fuente: Elaboración propia

Por otro lado, si se analizan los distintos tipos de ítems propuestos en cada uno de los años (Figura 2) se observa que no se proponen ejercicios relacionados con test o contrastes de hipótesis hasta el curso académico 2009/2010. A partir de entonces, casi la mitad de los ejercicios propuestos lo incluyen. Estos problemas parecen ir sustituyendo a los relacionados con el intervalo de confianza, que van disminuyendo, indicando un cambio de tendencia en los contenidos de inferencia evaluados en las pruebas de acceso. Este cambio podría afectar a la dificultad, pues autores como Cumming, Williams y Fidler (2004) indican que la comprensión de los intervalos es más sencilla que la de los contrastes tanto para estudiantes como para investigadores.

Figura 2: Clasificación por años de los tipos de problemas propuestos



Fuente: Elaboración propia

Finalmente resaltamos el hecho de que los problemas relacionados simplemente con muestreo (categorías A1 y A2) tienen una distribución muy variable. Aunque los encontramos todos los años, cabe destacar que a partir de 2010 solo han sido propuestos los ejercicios relacionados con la categoría A2 en 2011.

Conclusiones

Todas las pruebas de los doce años analizados incluyen un problema de inferencia estadística. Aunque el contenido de la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II incluye otros muchos temas, se destaca la importancia dada al aprendizaje de la inferencia (25% de la puntuación). Si se añade otro problema dedicado a la probabilidad (otro 25%) que hemos analizado en trabajos previos, el bloque de contenidos que recibe más atención en estas pruebas es el Bloque de Estadística y Probabilidad. Es paradójico el hecho de que tanto la estadística como la probabilidad sean, con frecuencia, los temas que se dejan en los cursos previos al Bachillerato como último tema y a veces se omite su enseñanza, mientras que luego el estudiante, que se prepara para selectividad, tendrá que hacer un esfuerzo notable para adquirir suficiente competencia y comprensión para resolver los problemas propuestos en las PAU y se juegue la mitad de su nota de esta prueba en mostrar la adquisición y desempeño de estos contenidos.

Dentro de los ejercicios relacionados con inferencia estadística, objetivamente se ve que las pruebas dan mayor importancia al cálculo intervalos de confianza y la interpretación de los mismos, pues aproximadamente el 64% de los ítems propuestos evalúan el aprendizaje de este tema, aunque, como hemos indicado, se observa un cambio de tendencia en los últimos años.

Analizando la solución de un ejemplo de cada categoría de campos de problemas propuestos (el resto de problemas propuestos en cada campo son similares al ejemplo analizado) se ha observado su complejidad. Ello se debe a que incluyen gran cantidad de objetos matemáticos, entre otros, población, muestra, parámetro poblacional, estadístico muestral, distribución del estadístico muestral. Se añaden los específicos del contraste de hipótesis e intervalo de confianza en los problemas relacionados con estos campos. En realidad, la resolución correcta de este problema no implica que el estudiante comprenda y discrimine todos estos conceptos, ni que haya adquirido suficiente razonamiento estadístico, sino que recuerda y sabe aplicar una serie de fórmulas, que quizás no comprenda. En resumen, nuestro análisis indica una alta dificultad de los problemas propuestos de inferencia estadística en las pruebas de acceso a la universidad, que debería ser tenida en cuenta por los diseñadores de las mismas en las sucesivas ediciones o en pruebas de evaluación alternativas que se propongan en el futuro.

Hacemos notar, finalmente, que esta prueba de evaluación presenta la inferencia, exclusivamente desde el punto de vista frecuencial, como una metodología única, ocultando las diferentes aproximaciones y las controversias que dentro de la misma estadística ha tenido la inferencia (BATANERO, 2000).

Agradecimientos: Proyecto EDU2013-41141-P (MEC) y grupo FQM126 (Junta de Andalucía).

Notas:
*Doctora en Matemáticas. Universidad de Granada/UGR. E-mail: mariadelmarlopez@ugr.es
**Doctora en Estadística. Universidad de Granada/UGR. E-mail:batanero@ugr.es
***Doctora en Psicología. Universidad de Huelva/UHU. E-mail:carmen.diaz@dpsi.uhu.es
****Doctora en Ciencias de la Educación. Universidad de Granada/UGR. E-mail:mmgea@ugr.es

Referencias

ARTEAGA, P. **Evaluación de conocimientos sobre gráficos estadísticos y conocimientos**

didácticos de futuros profesores. Tesis Doctoral (Doctorado en Didáctica de la Matemática). Universidad de Granada, Granada, 2011.

BATANERO, C. Controversies around significance tests. **Mathematical Thinking and Learning**, v. 2, n. 1, p. 75-98, 2000.

BISQUERRA, R. **Métodos de investigación educativa**. Barcelona: P.P.U., 1989.

BLÁZQUEZ, F.; LUENGO, R. **Las calificaciones en las pruebas de acceso en la Universidad de Extremadura**. Badajoz: ICE de la Universidad de Extremadura, 1989.

CARABALLO, R. **Análisis de los ítems de las pruebas de evaluación de diagnóstico en competencia matemática para el segundo curso de la educación secundaria obligatoria en España, 2008-2009: un estudio exploratorio**. 2010. 69 f. Trabajo fin de Máster (Máster en Didáctica de la Matemática) - Facultad de Ciencias de la Educación, Universidad de Granada, Granada, 2010.

CASTELLANOS, M. T. **Tablas y gráficos estadísticos en las pruebas Saber Colombia**. Trabajo fin de Máster (Máster en Didáctica de la Matemática) - Facultad de Ciencias de la Educación, Universidad de Granada, Granada, 2013.

CUMMING, G.; WILLIAMS, J.; FIDLER, F. Replication, and researchers' understanding of confidence intervals and standard error bars. **Understanding Statistics**, v. 3, p. 299-311, 2004.

DÍAZ, C.; CONTRERAS, J. M.; BATANERO, C.; ROA, R. Evaluación de sesgos en el razonamiento sobre probabilidad condicional en futuros profesores de educación secundaria. **Bolema**, v. 26, n. 22, p. 1207-1226, 2012.

ESPINEL, M. C.; RAMOS, C. E.; RAMOS. R. M. Identificación de los errores en los contrastes de hipótesis de los alumnos de Bachillerato. **SUMA**, v. 61, p. 35-44, 2009.

FALK, R.; GREENBAUM, C. W. Significance tests die hard: The amazing persistence of a probabilistic misconception. **Theory and Psychology**, v. 5, n. 1, p. 75-98, 1995.

GAVIRIA, J. L. La equiparación del expediente de Bachillerato en el proceso de selección de alumnos para el acceso a la universidad. **Revista de Educación**, v. 337, p. 351-387, 2005.

GODINO, J. D. Mathematical concepts, their meanings and understanding. En: PUIG, L.; GUTIÉRREZ, A. (Ed.), **Proceedings of the 20th PME Conference**, Valencia, Universidad de Valencia, v.2, p. 417-424, 1996.

GODINO, J. D.; BATANERO, C.; FONT, V. The onto-semiotic approach to research in mathematics education. **ZDM**, v. 39, n. 1, p. 27-135, 2007.

HARRADINE, A.; BATANERO, C.; ROSSMAN, A. Students and teachers' knowledge of sampling and inference. En BATANERO, C.; BURRILL, G.; READING, C. (Ed.), **Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education. Netherlands: Springer,** p. 235-246, 2011.

KRIPPENDORFF, K. **Metodología de análisis de contenido: teoría y práctica**. Barcelona: Paidos, 1997.

LÓPEZ-MARTÍN, M. M.; CONTRERAS, J. M.; BATANERO, C.; CARRETERO, M. Los problemas de probabilidad propuestos en las Pruebas de Acceso a la Universidad en Andalucía. **Areté**, v. 1, n. 1, p. 39-60, 2015.

MINGORANCE, C. **La estadística en las pruebas de diagnóstico andaluzas**. Trabajo fin de Grado (Grado de Maestro de Educación Primaria) - Facultad de Ciencias de la Educación, Universidad de Granada, Granada, 2014.

MEC, Ministerio de Educación y Ciencia. **Real Decreto 1467/2007, de 2 de noviembre 2007. Establece la estructura del Bachillerato y se fijan sus enseñanzas mínimas**. Boletín Oficial del Estado, Madrid, n. 266, 06 nov., 2007.

MECD, Ministerio de Educación, Cultura y Deporte. **Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato**. Boletín Oficial del Estado, Madrid, n. 3, 03 en., 2015.

MP, Ministerio de la Presidencia. **Real Decreto 1892/2008, de 14 de noviembre, por el que se regula las condiciones para el acceso a las enseñanzas universitarias oficiales de grado y los procedimientos de admisión a las universidades públicas españolas.** Boletín Oficial del Estado, Madrid, n. 138, 07 jun., 2008.

RICO, L. Marco teórico de evaluación en PISA sobre matemáticas y resolución de problemas. **Revista de Educación**, (Extraordinario 2006), p. 275-294, 2006.

RUIZ DE GAUNA GOROSTIZA, J.; DÁVILA BALSERA, P.; ETXEBERRIA MURGIONDO, J.; SARASUA FERNÁNDEZ, J. M. Pruebas de selectividad en Matemáticas en la UPV-EHU. Resultados y opiniones de los profesores. **Revista de Educación**, v. 362, p. 217-246, 2013.

SALDANHA. L.; THOMPSON, P. Conceptions of sample and their relationship to statistical inference. **Educational Studies in Mathematics**, v. 51, p. 257-270, 2002.

SCHUYTEN, G. Statistical thinking inpsychology and education. En D. VereJones (Ed.), Third International Conference on Teaching Statistics. **Proceedings** (pp. 486-490). Voorburg, The Netherlands, 1991.

VALLECILLOS, A. Some empirical evidence on learning difficulties about testing hypotheses. En 52 session of the International Statistical Institute**. Proceedings** (v.2, p.201–204). Helsinki, 1999.