

LOS INTERVALOS DE CONFIANZA EN LAS PRUEBAS DE SELECTIVIDAD ANDALUZAS

M^a del Mar López-Martín, Pedro Arteaga, María M. Gea, J. Miguel Contreras

Universidad de Granada. (España)

mariadelmarlopez@ugr.es, parteaga@ugr.es, mmgea@ugr.es, jmcontreras@ugr.es

RESUMEN: En este trabajo se analizan 144 pruebas de acceso a la universidad en Andalucía durante los últimos 12 años. De los tres bloques de contenidos que contiene la prueba, se han caracterizado los problemas sobre intervalos de confianza. Se observa un balance entre la construcción e interpretación de intervalos y la determinación del tamaño de muestra para una precisión o una amplitud de intervalo dados. Las poblaciones suelen ser binomiales o normales y el parámetro a estimar la media o la proporción. Se encuentra una proporción importante de ejercicios descontextualizados. Los resultados pueden ser tenidos en cuenta por los profesores a la hora determinar los temas en los que deben hacer énfasis con el fin de propiciar una mejor preparación de los estudiantes para asegurar el éxito en la prueba, que determina que el alumno pueda cursar la carrera deseada.

Palabras clave: pruebas PAU, bachillerato, intervalos confianza

ABSTRACT: This work deals with a study of 144 entrance examinations applied at the University of Andalucía during the last twelve years. From the three blocks of contents included in the proof, the problems about confidence intervals were characterized. As a result of this study, a balance was detected between the intervals' construction and interpretation, and the definition of the sample size, for the accuracy or amplitude of given intervals. The populations are usually binomial or normal, while the estimating parameter is the average or the proportion. There were a lot of non-contextualized exercises. The results can be taken into account by the teachers to define which topics they should focus on, in order to improve the students' training and to achieve the success in the entrance examination which determines the students' possibility to study the major they wish.

Key words: university entrance examination, high school, confidence intervals

■ Introducción

La importancia que dentro del aula se da a cada uno de los temas recogidos en las orientaciones curriculares viene indirectamente determinada por las pruebas de evaluación. Hoy en día la enseñanza de las matemáticas se ha visto fuertemente condicionada por los resultados que tienen los estudiantes en las pruebas durante las distintas etapas educativas. Ejemplo de ello es el impacto mediático que tienen las pruebas PISA (Program for International Student Assessment, OCDE, 2015), la evaluación TIMSS (Third International Mathematics and Science Study; Mullis, Martín, Foy y Arora, 2012) y las pruebas de diagnóstico y las Pruebas de Acceso a la Universidad (PAU).

Centrando el foco en la etapa de bachillerato (1º y 2º curso), es evidente la repercusión que tiene la calificación de las PAU, examen imprescindible para todos los estudiantes que desean continuar sus estudios en centros universitarios. El resultado de las pruebas PAU, junto con la calificación obtenida en bachillerato, determina la nota de admisión de acceso a la universidad y como consecuencia a los estudios deseados. Con el fin de asegurar el éxito de los futuros universitarios es necesario garantizar una adecuada correspondencia entre el contenido pretendido en las orientaciones curriculares para un tema y el evaluado en las mencionadas pruebas.

En el presente trabajo se fija como principal objetivo analizar el contenido de los ejercicios sobre intervalos de confianza (IC) que han sido recogidos en las pruebas PAU de la Comunidad de Andalucía en el periodo comprendido desde 2003 hasta 2014. Concretamente, se han analizado los problemas (consideramos el término problema en sentido muy amplio, como cualquier ejercicio o tarea que el alumno ha de resolver) correspondientes a IC que son incluidos en las pruebas que evalúan los contenidos tratados en la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II.

■ Fundamentos

En la actual sección se presentan los fundamentos teóricos sobre los que se sustenta el presente documento. En primer lugar, se analiza cómo tratan las normativas curriculares los contenidos relacionados con IC y las características principales de las pruebas PAU y, se finaliza mostrando un breve resumen del marco teórico que ha sido empleado.

Los intervalos de confianza en el currículo de Bachillerato y las pruebas PAU

A pesar de la importancia que tiene hoy en día el estudio de la inferencia estadística, solamente es posible encontrarla en el Bloque de Estadística y Probabilidad de la rama de conocimiento de Ciencias Sociales (MEC, 2007; MECD, 2015). La inclusión de la inferencia estadística en segundo curso de bachillerato tiene como objetivo forjar la base de algunos de los contenidos que una gran parte de los estudiantes seguirán tratando en la etapa universitaria e incluso en los ciclos formativos de la Formación Profesional.

El análisis de la normativa curricular vigente hasta el curso 2015/2016 revela que de los cinco puntos relacionados con inferencia estadística sólo uno está enfocado al estudio de los IC, donde se trata concretamente “Intervalos de confianza para el parámetro p de una distribución binomial y para la media de una distribución normal de desviación típica conocida” (MEC, 2007, p.45476). Sin embargo, la introducción de la Ley Orgánica para la Mejora de la Ley Educativa asigna un mayor protagonismo a los IC ya que la mitad de los contenidos fijados tratan sobre ellos (MECD, 2015, p.389).

Las actuales Pruebas de Acceso a la Universidad están regidas por el Real Decreto 1892/2008, por el que se regulan 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, en el artículo 38 de la Ley Orgánica de Educación (LOE), se pone de manifiesto que las pruebas PAU son pruebas de madurez que evalúan los conocimientos y capacidades adquiridas por los futuros universitarios durante su etapa de Bachillerato. La prueba correspondiente a Matemáticas Aplicadas a las Ciencias Sociales II 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 una de ellas. Ambas opciones están divididas en tres tipologías de ejercicios, un ejercicio correspondiente al Bloque Números y Álgebra, otro del Bloque de Análisis, y dos ejercicios del Bloque de Estadística y Probabilidad, concretamente, a Probabilidad e Inferencia Estadística. En la actualidad cada ejercicio tienen asignada una puntuación máxima de 2,5, por lo que el 50% de la nota total corresponde al Bloque de Estadística y Probabilidad.

■ Marco teórico

Nos apoyamos en algunos elementos del enfoque onto-semiótico (Godino, Batanero y Font, 2007), que diferencia entre el significado institucional de un objeto matemático (compartido dentro de una institución) y el significado personal del mismo (que adquiere una persona). Se centra la atención en la institución de enseñanza, donde se diferencia el significado global (en nuestro caso, el significado del intervalo de confianza), el significado pretendido dentro de un nivel educativo (en nuestro caso, el marcado por las orientaciones curriculares) y el evaluado (contenido en las pruebas de evaluación). El presente estudio se orienta a comparar los significados pretendido y evaluado del intervalo de confianza en las PAU de Matemáticas Aplicadas a las Ciencias Sociales II. La importancia de este punto se fundamenta en que las respuestas a las pruebas de evaluación dan acceso a la comprensión del tema por parte del estudiante.

Puesto que la comprensión de un determinado objeto matemático es un constructo psicológico inobservable, Godino (1996) estableció que el análisis de las respuestas de los estudiantes a los ítems, tareas o pruebas de evaluación, puede permitir obtener cierta información sobre la comprensión que tiene el estudiante en el tema de estudio. Por tanto, dado el vínculo existente entre el significado institucional pretendido y el evaluado para un contenido, es fundamental que esa correspondencia sea idónea. Para tal fin, será primordial garantizar que las pruebas de evaluación sean coherentes con los contenidos fijados institucionalmente; así pues, en esta investigación se trata de evaluar la adecuación

de esa correspondencia para el contenido pretendido de los IC en las orientaciones curriculares de Bachillerato.

Asimismo, en la elaboración del presente trabajo se han considerado investigaciones llevadas en las que se analizan las PAU desde otros puntos de vista. Por ejemplo, Ruiz, Dávila, Etxeberria y Sarasua (2013) analizan los resultados de los estudiantes y la opinión del profesorado, mientras que Belia, Fidler y Cumming (2005) y Olivo (2008) se centran más en describir la comprensión que tienen tanto los estudiantes como investigadores sobre IC. Estos últimos trabajos han sido empleados con el fin de elegir las variables del estudio que aquí se presenta.

■ Metodología

Mediante un análisis de contenido, cuyo foco de atención se centra en el análisis del significado del concepto a través de la descripción de los significados y la correcta aplicación del contenido, se han clasificado los ejercicios de inferencia estadística propuestos de acuerdo a los siguientes campos de problemas:

- Composición de muestras y cálculo de estadísticos de las distribuciones muestrales en poblaciones finitas.
- Distribución muestral en poblaciones infinitas y cálculo de probabilidades.
- Cálculo y/o interpretación de un Intervalo de Confianza.
- Relación entre confianza, error de estimación y tamaño muestral.
- Contraste de hipótesis.

Del conjunto total de ejercicios de inferencia estadística, se ha centrado la atención en los que tratan los contenidos de IC. Igualmente, de las actividades seleccionadas se ha realizado un estudio de la distribución de las siguientes variables:

- El modelo de distribución de la población de partida en el ejercicio.
- El parámetro que se pide estimar.
- El contexto en el que se presenta la tarea, dentro de los considerados en las pruebas PISA (OCDE. 2009).

Puesto que la metodología empleada pretende dar respuesta a problemas concretos, la selección de la muestra se ha realizado de forma intencionada. Así pues, es importante señalar que el objetivo de esta investigación no es extrapolar los resultados a otras pruebas diferentes a las analizadas. No obstante, con el presente estudio se pretende aportar algunas ideas que pueden servir para conocer el

contenido de inferencia estadística de las PAU. Asimismo, los resultados obtenidos 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 útiles para los especialistas universitarios que forman parte de las Comisiones Coordinadoras de las PAU.

Muestra de problemas

Cada año la Comunidad Autónoma de Andalucía convoca en los meses de junio y septiembre las pruebas PAU. Junto a las pruebas realizadas en las convocatorias oficiales se proponen cuatro exámenes de reserva, por lo que hay un total seis modelos disponibles por año. Dado que cada modelo contiene dos ejercicios de inferencia estadística (uno por cada opción), se analizaron 144 problemas. Del total de problemas seleccionados, se comprobó que algunos de ellos abarcaban más de un campo de problema, por lo que el número total de apartados analizados fueron 285, de los que más de la mitad correspondían a IC (175). Este hecho permite deducir la gran importancia que tienen los IC en las pruebas PAU.

Campos de problemas

Los ítems relacionados con IC se han clasificado en dos campos de problemas que se presentan con similar frecuencia y que se detallan a continuación.

P1. *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 parámetro (generalmente la media o la proporción poblacional), así como los relacionados con la interpretación de la información aportada por el intervalo, siendo este último punto más difícil, según Belia, Fidler y Cumming (2005) y Olivo (2008).

Recordemos que la construcción de un intervalo de confianza depende de:

- Las características de la población, es decir, si el modelo probabilístico es binomial, normal, etc.
- Los parámetros poblacionales desconocidos, como pueden ser por ejemplo la media, varianza y proporción.
- Los valores de los estadísticos muestrales conocidos.
- Tamaño muestral.

Aunque a la hora de determinar un intervalo hay que tener en cuenta diversos factores, si bien es cierto, la forma de actuar en su construcción permanece invariable. No obstante, aunque la

construcción del intervalo no presenta grandes dificultades para los estudiantes es necesario señalar que la interpretación no es algo tan obvio (Cumming y Fidler, 2005; Harradine, Batanero y Rossman, 2011). Una vez determinado el intervalo, no es correcto interpretar que el nivel de confianza es la probabilidad de que el verdadero valor del parámetro esté contenido en el intervalo, ya que, una vez definidos los valores de los extremos, se pierde el carácter de aleatoriedad y por tanto, la probabilidad será 1 si el intervalo es uno de los que contienen al parámetro, o 0 si es uno de los restantes intervalos que no lo contienen. Por esta razón, cuando se tiene un valor concreto del intervalo de confianza se habla de confianza y no de probabilidad, es decir, confiamos en que el intervalo calculado sea el que contiene el verdadero valor (Cumming, Williams y Fidler, 2004).

P2. Relación entre confianza, error de estimación y tamaño muestral. En la práctica, cuando se pretende obtener un intervalo de confianza para un parámetro poblacional siempre se intenta describir aquel intervalo que conduce al menor error de estimación posible. Puesto que, en general, la expresión del intervalo de confianza depende de la estimación del parámetro, del tamaño muestral y del nivel de confianza seleccionado, será necesario determinar la relación existente entre ellos con el fin de obtener una estimación lo más precisa posible. A partir de conexión existente entre los tres elementos se pueden establecer las siguientes propiedades (Behar, 2001):

- 1) A medida que aumenta el tamaño de la muestra (manteniendo constante el nivel de confianza) el error de estimación disminuirá.
- 2) Manteniendo constante el tamaño de la muestra, el error aumentará conforme aumente el nivel de significación.

No obstante, es necesario señalar que si el objetivo es conseguir una mayor precisión en las estimaciones realizadas por medio de los IC, lo ideal es aumentar el tamaño muestral, ya que habitualmente no es recomendable utilizar un nivel de confianza inferior al 90%.

En este bloque se han analizado aquellos ítems en los que se solicita calcular el nivel de confianza, error de estimación o tamaño muestral después de haber fijado dos de ellos. Como datos recogidos en el enunciado se suele encontrar la población y el estadístico considerado o bien el intervalo de confianza.

■ Resultados

A partir de la identificación de los distintos campos de problemas se han encontrado que de los 285 apartados, 86 corresponden a problemas de tipo P1 y el resto (89) a P2. El estudio longitudinal en el periodo seleccionado muestra una distribución similar entre ambos campos de problemas (véase Figura 1).

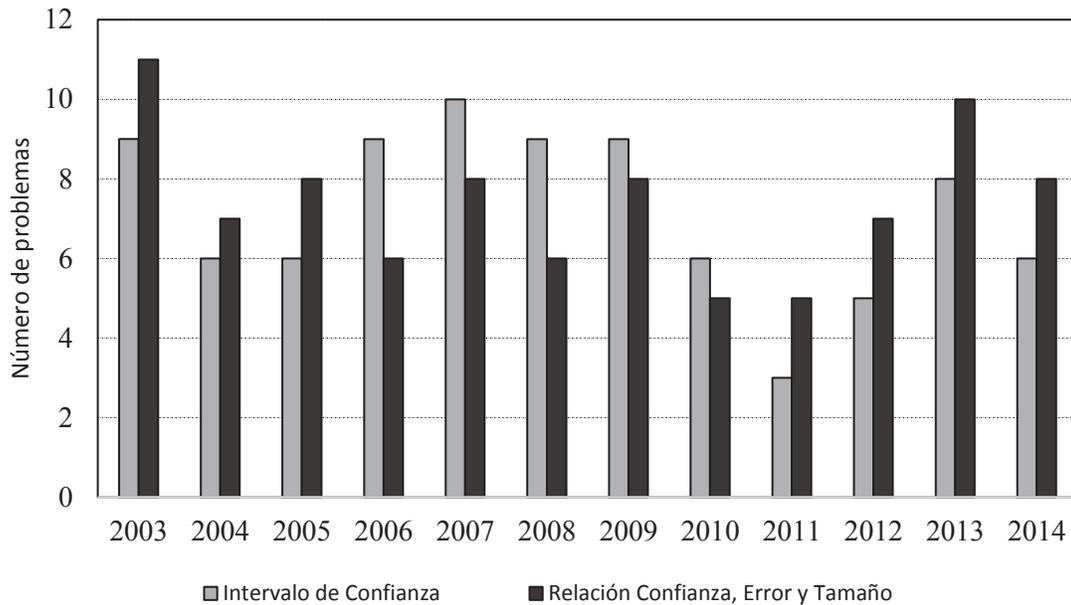


Figura 1. Clasificación por años según los campos de problemas

A consecuencia de que el proceso de construcción de un intervalo de confianza parte del hecho de que la población de partida se distribuye según un modelo teórico determinado, se identificó el modelo probabilístico y el parámetro objeto de estudio (véase Tabla 1).

Del análisis se determina que las distribuciones de probabilidad empleadas son: la distribución binomial (36%) y la distribución normal con desviación típica conocida (dato facilitado en el enunciado del problema) (64%). Puesto que la construcción del intervalo de confianza depende del estadístico muestral y de la distribución que siga éste, y teniendo en cuenta que la proporción muestral se distribuye según una binomial y que la media muestral sigue una distribución normal, los resultados obtenidos en el análisis del parámetro a estimar coinciden con los del modelo probabilístico de la población de partida.

Tabla 1. Modelo probabilístico y parámetro a estimar

Modelo probabilístico	Parámetro a estimar	
	Proporción	Media
Distribución Binomial	36%	
Distribución Normal		64%

Se destaca que, aunque el uso de la distribución binomial es menor que el de la distribución normal, el empleo de ella incrementa la dificultad en cuanto que requiere utilizar la aproximación a la distribución normal para la construcción del intervalo.

■ Síntesis del análisis

El análisis de los problemas sobre IC incluidos en las pruebas PAU pone de manifiesto la necesidad de dominar una gran variedad de conceptos y procedimientos. En ambos problemas se han observado conocimientos comunes evaluados, en particular los de: Población y muestra; Distribución muestral; Relación entre los estadísticos de la población y de la distribución muestral; Modelo teórico de distribución; Distribución Normal tipificada. Ello fuerza al estudiante a diferenciar los tres planos de distribución señalados por Schuyten (1991):

- La distribución de los datos en la muestra.
- La distribución de la variable de interés en la población.
- La distribución muestral del estadístico en todas las posibles muestras obtenidas de la población.

Otros conceptos y propiedades evaluadas en ambos problemas han sido: Estimador de un parámetro; Intervalo de confianza; Coeficiente de confianza o nivel de confianza; Probabilidad condicional. Como diferencia se señala que los conceptos de relación entre amplitud de intervalo, precisión y tamaño muestral solamente se evalúan únicamente con los problemas P2.

En cuanto a los procedimientos que se llevan a cabo en la realización de los distintos problemas, se destaca la importancia que tiene el cálculo de los estadísticos en las muestras con el fin de obtener las

medidas de centralización y proporción muestral para posteriormente describir la distribución tanto de la media muestral como de la proporción muestral. Otro punto importante a tener en cuenta es una lectura correcta.

Puesto que las pruebas PAU tienen como objetivo evaluar los conocimientos y capacidades adquiridas por los estudiantes en la etapa de bachillerato, es necesario asegurar que existe una adecuada correspondencia entre los contenidos incluidos en el currículo de bachillerato y los evaluados en las pruebas. El análisis de los contenidos fijados en las directrices curriculares ha permitido comprobar que la mitad de los criterios de evaluación sobre inferencia son evaluados con los problemas de IC.

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. Además, permiten efectuar recomendaciones para mejorar la enseñanza, introduciendo aproximaciones informales que permitan a los estudiantes adquirir el sentido del intervalo de confianza.

Agradecimientos: Proyectos EDU2013-41141-P y EDU2016-74848-P (AEI, FEDER) y Grupo FQM126 (Junta de Andalucía)

■ Referencias Bibliográficas

- Behar, R. (2001). *Aportaciones para la mejora del proceso de enseñanza-aprendizaje de la estadística*. Tesis Doctoral. Universidad Politécnica de Cataluña.
- Belia, S., Fidler, F. y Cumming, G. (2005). Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods*, 4, 389-396.
- Cumming, G. y Fidler, F. (2005). Interval estimates for statistical communication: problems and possible solutions, IASE/ISISatellite. Disponible en: www.stat.auckland.ac.nz/~iase/publications/14/cumming.pdf
- Cumming, G., Williams, J. y Fidler, F. (2004). Replication and researchers' understanding of confidence intervals and standard error bars. *Understanding statistics*, 3(4), 299-311.
- 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, p. 417-424). Universidad de Valencia.
- 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), 27-135.
- 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 Netherlands.

- MEC, Ministerio de Educación y Ciencia. (2007). *Real Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura del Bachillerato y se fijan sus enseñanzas mínimas*. Madrid: Autor.
- MECD, Ministerio de Educación, Cultura y Deporte. (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato*. Madrid: Autor.
- MP, Ministerio de la Presidencia (2008). *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*. Madrid: Autor.
- Mullis, I. V., Martin, M. O., Foy, P. y Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Amsterdam: International Association for the Evaluation of Educational Achievement.
- OCDE. (2009). *PISA 2009 assessment framework - key competencies in reading, mathematics and science*. Paris: OCDE.
- Olivo, E. (2008). *Significados del intervalo de confianza en la enseñanza de la ingeniería en México*. Tesis Doctoral. Universidad de Granada.
- Ruiz, J. G., Dávila, P., Etxebarria, J. y Sarasua, J. (2013). Pruebas de selectividad en Matemáticas en la UPV-EHU. Resultados y opiniones de los profesores. *Revista de educación*, 362, 217-246.
- Schuyten, G. (1991). Statistical thinking in psychology and education. En D. VereJones (Ed.), *Proceeding of the Third International Conference on Teaching Statistics* (pp. 486-490). Voorburg, The Netherlands: International Statistical Institute.