

**VARIABLES QUE DETERMINAN LOS PROBLEMAS  
DE INFERENCIA PROPUESTOS EN ANDALUCÍA EN  
LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD**

María del Mar López-Martín



TRABAJO FIN DE MÁSTER

Tutores: Dra. Carmen Batanero Bernabeu y Dr. Rafael Roa Guzmán

Universidad de Granada  
Departamento de Didáctica de la Matemática

Septiembre 2016



## AGRADECIMIENTOS

Quiero dejar mi más profundo agradecimiento a mis tutores Carmen Batanero Bernabeu y Rafael Roa Guzmán por su constante asesoramiento, implicación, tiempo y esfuerzo y por la confianza que han depositado en mí en todo momento.

Gracias a todos mis compañeros por su ayuda y amabilidad que me han mostrado desde el primer día.

Por último, agradecer a mi familia la paciencia y cariño aportado durante este periodo. En especial, a mi marido y mi pequeña por su apoyo condicional y su comprensión por todo el tiempo que le he robado para poder llevar a cabo este trabajo.

Trabajo realizado en el marco de los Proyectos EDU 2013-41141-P y EDU2016-74848-P (MEC) dentro del grupo de investigación FQM126 Teoría de la Educación Matemática y Educación Estadística (Junta de Andalucía).



<b>INTRODUCCIÓN.....</b>	<b>1</b>
<b>PLANTEAMIENTO DEL PROBLEMA.....</b>	<b>3</b>
1.1. INTRODUCCIÓN .....	3
1.2. LA INFERENCIA EN EL CURRÍCULO DE BACHILLERATO .....	4
1.3. LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD (PAU) .....	8
1.3.1. Las actuales pruebas PAU .....	10
1.3.2. El proyecto en marcha sobre la “nueva selectividad” .....	12
1.4. IMPORTANCIA DE LA INFERENCIA ESTADÍSTICA.....	13
1.4.1. Papel en el desarrollo de las ciencias.....	14
1.4.2. Importancia en la formación del estudiante.....	17
1.5. OBJETIVOS E HIPÓTESIS DE LA INVESTIGACIÓN .....	18
<b>FUNDAMENTOS.....</b>	<b>21</b>
2.1. INTRODUCCIÓN .....	21
2.2. MARCO TEÓRICO .....	22
2.2.1. Significados institucionales y personales y sus tipos .....	23
2.2.2. Evaluación en el EOS .....	24
2.2.3. Idoneidad didáctica y sus componentes.....	25
2.3. INVESTIGACIONES SOBRE LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD.....	27
2.3.1. Análisis de las pruebas de matemáticas.....	27
2.3.2. Estudio de la inferencia en las PAU .....	30
2.4. ANÁLISIS SOBRE OTRAS PRUEBAS DE EVALUACIÓN.....	31
2.5. DIFICULTADES DE LOS ESTUDIANTES CON LA INFERENCIA.....	32

2.5.1. Comprensión de la distribución muestral .....	33
2.5.2. Comprensión del contraste de hipótesis .....	35
2.5.3. Comprensión del intervalo de confianza .....	37
2.6. CONCLUSIONES .....	37

## **ANÁLISIS DE EJERCICIOS DE INFERENCIA ESTADÍSTICA EN LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD ..... 39**

3.1. INTRODUCCIÓN .....	39
3.2. METODOLOGÍA .....	41
3.3. MUESTRA DE PROBLEMAS ANALIZADOS .....	42
3.4. VARIABLES ANALIZADAS.....	43
3.5. CAMPOS DE PROBLEMAS Y CONTENIDO EVALUADO EN CADA UNO DE ELLOS.....	45
3.5.1. Análisis semiótico del contenido evaluado .....	46
3.5.2. Síntesis del análisis de los campos de problemas.....	59
3.5.3. Estudio comparativo entre el significado pretendido y evaluado.....	61
3.6. MODELO PROBABILÍSTICO DE LA POBLACIÓN DE PARTIDA .....	64
3.7. PARÁMETRO POBLACIONAL OBJETO DE ESTUDIO .....	66
3.8. CONTEXTO DEL PROBLEMA .....	68
3.9. ANÁLISIS ESTADÍSTICO DE LOS PROBLEMAS.....	70
3.9.1. Campos de problemas.....	70
3.9.2. Modelo probabilístico de la población de partida .....	73
3.9.3. Parámetro poblacional objeto de estudio.....	74
3.9.4. Contexto .....	76
3. 10. CONCLUSIONES SOBRE LA IDONEIDAD DIDACTICA DE LAS PAU .	77

## **CONCLUSIONES ..... 81**

4.1 CONCLUSIONES RESPECTO A LOS OBJETIVOS DEL TRABAJO E HIPÓTESIS.....	81
--	----

4.2. LIMITACIONES DEL ESTUDIO Y FUTURAS LÍNEAS DE INVESTIGACIÓN.....	85
4.3. IMPLICACIONES PARA LA ENSEÑANZA .....	85
<b>REFERENCIAS .....</b>	<b>87</b>
<b>ANEXO.....</b>	<b>97</b>
Trabajos que desarrollan esta memoria .....	97



# INTRODUCCIÓN

La enseñanza matemática se ha visto fuertemente condicionada por las pruebas de evaluación externas, pues es el deseo, tanto de los mismos alumnos, como de los padres y profesores, que cada centro obtenga buenos resultados en ellas. Dada su impacto mediático, destacamos, en este sentido, la influencia del proyecto PISA (Program for International Student Assessment; OCDE, 2015), la evaluación TIMSS (Third International Mathematics and Science Study; Mullis, Martín, Foy y Arora, 2012), las pruebas de diagnóstico y las Pruebas de Acceso a la Universidad (PAU).

Los resultados ofrecidos tanto por las pruebas PISA como por las pruebas TIMSS sugieren que, para muchos estudiantes, las matemáticas solamente son un conjunto de reglas y técnicas sin significado y sin relación con la vida cotidiana. La situación de fracaso que evidencian para estos alumnos los resultados de las pruebas nos hace reflexionar e investigar sobre los factores que pueden ser causa de ese fracaso y, en nuestro caso, nos interesa saber si en la etapa de Bachillerato se mantiene esa tendencia.

Puesto que el objetivo principal de las pruebas PAU es evaluar el rendimiento que han adquirido los estudiantes cuando finalizan sus estudios de Bachillerato, el análisis y estudio de las mismas nos aportará una gran información sobre la adecuación de los contenidos de estas pruebas. Así que, el principal objetivo que nos marcamos con la elaboración de este trabajo es analizar el contenido de los ejercicios de inferencia estadística propuestos en las pruebas PAU del Distrito Andaluz. Concretamente, se han seleccionado las pruebas de la especialidad de Bachillerato de Ciencias Sociales propuestas en el período desde 2003 a 2014. Se trata de clasificar estos ejercicios respecto a ciertas variables de interés y analizar la distribución de los mismos a lo largo del periodo. Posterior al análisis nos planteamos como objetivo comprobar si el contenido de los problemas relacionados con inferencia estadística reflejan adecuadamente las recomendaciones recogidas en las orientaciones curriculares nacionales.

La memoria se ha estructurado en cuatro capítulos. En el primero de ellos se realiza el planteamiento del problema analizando cómo el currículo de Bachillerato aborda los contenidos de inferencia estadística. Junto a una revisión sobre los cambios que ha sufrido las PAU desde sus inicios, se han presentado las características de las pruebas

vigentes hasta el actual curso académico y se muestran algunas pinceladas sobre la revalida que sustituirá el próximo año académico a la actual selectividad. Para finalizar se justifica la importancia que tiene el estudio de la inferencia estadística y se describen los objetivos sobre los que se ha desarrollado la investigación e hipótesis o expectativas iniciales que se esperan encontrar en la elaboración del trabajo.

En el segundo capítulo se presenta el marco teórico que sustenta la investigación realizada, junto con una síntesis de las investigaciones sobre el análisis de pruebas que se han llevado a cabo por otros autores. Con objeto de intentar analizar el porqué de los fracasos que se producen, se realiza una revisión bibliográfica sobre los trabajos de investigación centrados en el estudio de las dificultades que presentan los estudiantes cuando tratan los contenidos de inferencia.

En el tercer capítulo, mediante un análisis de contenido se han clasificado los ejercicios propuestos de acuerdo a diferentes campos de problema diferenciados en inferencia. Asimismo, se ha analizado el modelo de distribución de la población de partida en el ejercicio; el parámetro que se pide estimar y el contexto en el que se presenta el ejercicio, dentro de los considerados en las pruebas PISA. Para cada una de las variables se ha realizado un estudio de la distribución de las diferentes categorías, globalmente y a lo largo del periodo seleccionado. A partir del análisis realizado, se lleva a cabo una reflexión sobre algunos componentes de la idoneidad didáctica de las pruebas PAU.

Para finalizar la memoria, el último capítulo recoge las conclusiones más destacadas del trabajo. Dada las limitaciones de espacio se incorpora en un anexo el listado de los trabajos de investigación frutos de esta memoria que han sido publicados en congresos internacionales y en revistas de catálogo Latindex.

# Capítulo 1

## PLANTEAMIENTO DEL PROBLEMA

### Índice de Contenidos

---

- 1.1. Introducción
  - 1.2. La inferencia en el currículo de Bachillerato
  - 1.3. Las Pruebas de Acceso a la Universidad
  - 1.4. Importancia de la inferencia estadística
  - 1.5. Objetivos e hipótesis de la investigación
- 

### 1.1. INTRODUCCIÓN

La importancia que dentro del aula se da a cada tema recogido en las orientaciones curriculares viene indirectamente determinada por las pruebas de evaluación, en particular, para los estudiantes de Bachillerato, por las Pruebas de Acceso a la Universidad. El artículo 7 del Real Decreto 1892/2008, de 14 de noviembre, 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), establece que:

1. La prueba se adecuará al currículo del Bachillerato y versará sobre las materias, a las que se refieren los artículos 6 y 7 del Real Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura del Bachillerato y se fijan sus enseñanzas mínimas, establecidas para el segundo curso.
2. Las Administraciones educativas y las universidades organizarán la prueba y deberán garantizar su adecuación al currículo del Bachillerato, así como la coordinación entre las universidades y los centros que imparten Bachillerato para la organización de la misma. (p. 46934).

Aunque inicialmente la prueba está enfocada para seleccionar los estudiantes que desean ingresar en determinados estudios y centros universitarios, es evidente que su diseño está más orientado en la evaluación de los conocimientos y capacidades adquiridas por los futuros universitarios durante su etapa de Bachillerato.

En este primer capítulo se presenta con mayor precisión el problema que está siendo objeto de estudio. Dada a la estrecha relación que existente entre Bachillerato y las Pruebas de Acceso a la Universidad (PAU), se realiza en primer lugar una revisión de cómo las orientaciones curriculares de Bachillerato tratan el tema de inferencia estadística, para posteriormente analizar la evolución, normativa, estructura y contenido de las pruebas PAU. Como en la actualidad nos encontramos en un escenario, algo inestable e impreciso, relacionado con la eliminación de las pruebas PAU y la incorporación de una nueva reválida, se ha visto conveniente incluir en la Sección 1.3 la información, con la que se puede contar hoy en día, en relación a la nueva propuesta.

Dado que el contenido que se desea evaluar con la presentación del Trabajo Fin de Máster es el de inferencia estadística, en la Sección 1.4 se señala el alcance que ha adquirido en las últimas décadas la enseñanza de la inferencia así como la importancia que tiene en la formación de los estudiantes. Para finalizar, se describen los objetivos que sustentan el presente trabajo y se formulan algunas hipótesis, que se entiende cómo expectativas iniciales.

## **1.2. LA INFERENCIA EN EL CURRÍCULO DE BACHILLERATO**

A pesar de la importancia que tiene hoy en día la inferencia estadística, en el Real Decreto 1467/2007, de 2 de noviembre, donde se establece la estructura del Bachillerato y se fijan sus enseñanzas mínimas (MEC, 2007), solamente es posible encontrar los contenidos de inferencia estadística en el Bloque de Estadística y Probabilidad incluido en la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II. Este contenido se omite para los estudiantes que cursan el Bachillerato de Ciencias y Tecnología (Ciencias en la LOMCE), que también necesitan conocimientos sobre inferencia en sus futuros estudios.

La inclusión de la inferencia estadística en segundo curso de Bachillerato está enfocada en formar la base del andamiaje de algunos de los contenidos que serán tratados posteriormente en la etapa universitaria e incluso en los ciclos formativos de la Formación Profesional.

Según el Real Decreto (MEC, 2007), los contenidos fijados para el tema de inferencia estadística son los siguientes:

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. (p. 45476).

Si se compararan los contenidos incluidos en las orientaciones curriculares con los que se pueden encontrar en los primeros cursos de grado, se observaría que éstos abarcan prácticamente un curso elemental de inferencia estadística. El hecho de que los estudiantes que cursan el Bachillerato de Humanidades y Ciencias Sociales reciben menor formación matemática, que los del Bachillerato de Ciencias y Tecnología (Ciencias en la LOMCE), sobre contenidos de álgebra lineal o el análisis matemático, junto con el escaso tiempo que disponen para trabajar el tema, dificulta una comprensión formal de la inferencia. Recordemos que la asignatura de Matemáticas Aplicada a las Ciencias Sociales II incluye, además del bloque de estadística y probabilidad, los bloques de números y álgebra, bloque de análisis y un bloque de contenidos transversal, relacionado con la resolución de problemas.

La situación política que estamos viviendo hoy en día está llena de incertidumbre y el tema de las leyes educativas lo es aún más. La actual Ley Orgánica para la Mejora de la Ley Educativa (LOMCE) es la séptima ley orgánica de la democracia española. La LOMCE fue llevada adelante por el Partido Popular sin el apoyo de ningún grupo parlamentario y hoy en día pende de un hilo. Desde su publicación, la implantación de la ley ha encontrado muchos obstáculos, ya que algunas comunidades autónomas mostraban su rechazo a la implantación de la reforma educativa.

En la comunidad andaluza y en base al Decreto 97/2015, de 3 de marzo, se establece la ordenación y las enseñanzas correspondientes a la educación primaria en Andalucía, implantando todos los cursos de educación primaria (CECD, 2015). En el caso de la Educación Secundaria Obligatoria y del Bachillerato la situación es totalmente diferente. Puesto que el Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de ambas etapas, se hizo público el día 3 de enero de 2015, no fue posible elaborar y aprobar la normativa con los nuevos aspectos de la LOMCE en la comunidad de Andalucía. Por ello, se decidió que el curso académico 2015/2016 fuera un curso transitorio, mientras se aprobaba la normativa autonómica de adaptación a la LOMCE. Durante el mencionado curso académico se implantó únicamente los cursos impares de Educación Secundaria Obligatoria (1º y 3º ESO) y de Bachillerato (1º Bachillerato). Para el próximo curso académico, 2016/2017, la situación no ha variado mucho, por lo que será de nuevo un curso de transición en el que la nueva ley se irá incorporando de forma progresiva en los cursos restantes.

Al igual que la LOE, la LOMCE incluye únicamente los contenidos de inferencia estadística en segundo curso de Bachillerato de la modalidad de Ciencias Sociales. Concretamente el Real Decreto 1105/2014, de 26 de diciembre, dentro de la asignatura Matemática II Aplicadas a las Ciencias Sociales, incluye los siguientes contenidos (MEC, 2015b):

Población y muestra. Métodos de selección de una muestra. Tamaño y representatividad de una muestra.

Estadística paramétrica. Parámetros de una población y estadísticos obtenidos a partir de una muestra. Estimación puntual.

Media y desviación típica de la media muestral y de la proporción muestral. Distribución de la media muestral en una población normal. Distribución de la media muestral y de la proporción muestral en el caso de muestras grandes.

Estimación por intervalos de confianza. Relación entre confianza, error y tamaño muestral.

Intervalo de confianza para la media poblacional de una distribución normal con desviación típica conocida.

Intervalo de confianza para la media poblacional de una distribución de modelo desconocido y para la proporción en el caso de muestras grandes. (p.389).

Si comparamos los contenidos incluidos en LOE y LOMCE observamos que, aunque un alto porcentaje de los contenidos son similares entre sí, en el caso de la LOMCE se especifican de una forma más detallada. Por ejemplo, en el caso de los intervalos de confianza se observa que, además de incluir la construcción del propio intervalo, la LOMCE señala la importancia de estudiar la relación entre confianza, error

y tamaño muestral. Aun así, la principal diferencia entre ambas leyes la encontramos con los contrastes de hipótesis que son eliminados en la LOMCE.

Como novedad en la nueva regulación (MECD, 2015b) se observa la inclusión, junto a los criterios de evaluación (criterios para evaluar, tanto en conocimientos como en competencias, el aprendizaje del alumnado), los estándares de aprendizaje evaluables en todas las asignaturas, que se definen como

Especificaciones de los criterios de evaluación que permiten definir los resultados de aprendizaje, y que concretan lo que el estudiante debe saber, comprender y saber hacer en cada asignatura; deben ser observables, medibles y evaluables y permitir graduar el rendimiento o logro alcanzado. Su diseño debe contribuir y facilitar el diseño de pruebas estandarizadas y comparables. (p. 172).

Puesto que los criterios de evaluación como los estándares de aprendizaje evaluables son esenciales en la planificación del currículo, en el Cuadro 1.1 se muestran los relacionados con el tema de inferencia estadística.

**Cuadro 1.1.** Criterios de evaluación y estándares de aprendizaje evaluables. (MECD, 2015b, p. 389)

Criterios de evaluación	Estándares de aprendizaje evaluables
Describir procedimientos estadísticos que permiten estimar parámetros desconocidos de una población con una fiabilidad o un error prefijados, calculando el tamaño muestral necesario y construyendo el intervalo de confianza para la media de una población normal con desviación típica conocida y para la media y proporción poblacional cuando el tamaño muestral es suficientemente grande.	<p>2.1. Valora la representatividad de una muestra a partir de su proceso de selección.</p> <p>2.2. Calcula estimadores puntuales para la media, varianza, desviación típica y proporción poblacionales, y lo aplica a problemas reales.</p> <p>2.3. Calcula probabilidades asociadas a la distribución de la media muestral y de la proporción muestral, aproximándolas por la distribución normal de parámetros adecuados a cada situación, y lo aplica a problemas de situaciones reales.</p> <p>2.4. Construye, en contextos reales, un intervalo de confianza para la media poblacional de una distribución normal con desviación típica conocida.</p> <p>2.5. Construye, en contextos reales, un intervalo de confianza para la media poblacional y para la proporción en el caso de muestras grandes.</p> <p>2.6. Relaciona el error y la confianza de un intervalo de confianza con el tamaño muestral y calcula cada uno de estos tres elementos conocidos los otros dos y lo aplica en situaciones reales.</p> <p>3.1. Utiliza las herramientas necesarias para estimar parámetros desconocidos de una población y presentar las inferencias obtenidas mediante un vocabulario y representaciones adecuadas.</p> <p>3.2. Identifica y analiza los elementos de una ficha técnica en un estudio estadístico sencillo.</p> <p>3.3. Analiza de forma crítica y argumentada información estadística presente en los medios de comunicación y otros ámbitos de la vida cotidiana.</p>

### 1.3. LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD (PAU)

Para estudiar la implantación de las Pruebas de Acceso a la Universidad tenemos que remontarnos a la década de los cuarenta (Muñoz Vitoria, 1995). En el trabajo presentado por Muñoz Vitoria (1995) sobre una perspectiva histórica de las Pruebas de Acceso a la Universidad, el autor distingue tres etapas según la denominación y el enfoque que adquiere la prueba.

- Primera etapa, desde 1940 hasta 1953. Por la Ley de Reforma de la segunda enseñanza del 20 de septiembre de 1938, se instaura el Examen de Estado. La realización del examen se lleva a cabo al finalizar el séptimo curso.
- Segunda etapa, desde 1953 hasta 1971. La prueba recibe el nombre de Prueba de Madurez y se implanta por la Ley sobre ordenación de la Enseñanza Media del 26 de febrero de 1953. La prueba de madurez la realizaban los estudiantes que habían cursado el conocido como Curso Preuniversitario.
- Tercera etapa, desde 1971 hasta 1990<sup>1</sup>. En 1970 se publica la Ley General de Educación y Financiamiento de la Reforma Educativa. Sin embargo, a lo largo de este período el cambio más significativo, con respecto a la prueba, se tiene en 1975. Con la Ley 30/1974 se instaura lo que se conoce como las Pruebas de Aptitud para el Acceso a la Universidad (PAAU) cuyo diseño estaba enfocado en reconocer a todos aquellos alumnos que están capacitados para iniciar dichos estudios (universitarios). En esta época la prueba se enfoca a los estudiantes que realizan el Curso de Orientación Universitaria (COU).

Las Pruebas de Aptitud para el Acceso a la Universidad han sido uno de los formatos más estables, manteniéndose hasta 1990. Durante ese año, se promulga la Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo (LOGSE) derogando la Ley General de Educación de 1970. La implantación de la LOGSE provocó un replanteamiento de las Pruebas de Acceso a la Universidad ya que se suprime el curso de orientación y preparación para el acceso a la Universidad (conocido como COU) y las directrices específicas para preparar a los alumnos, del segundo ciclo de enseñanza Secundaria, que deseaban ingresar en la universidad.

---

<sup>1</sup> En los periodos 1971-1972 y 1973-1974 se eliminaron las pruebas de acceso.

Analizando el comportamiento que ha tenido durante las dos últimas décadas, es fácil comprobar que esa supuesta estabilidad ha desaparecido pasando a un escenario de gran inestabilidad. Ejemplo de ello se muestra en la Figura 1.1 que incluye algunas de las modificaciones que se han realizado desde 1999 desde el punto de vista de la legislación. Desde 1999 y hasta la actualidad, en la página oficial del Ministerio de Educación, Ciencia y Deporte es posible encontrar seis reales decretos (dos de ellas son modificaciones sobre el Real Decreto 1640/1999 y 990/2000) cuyo objetivo son la regulación de las Pruebas de Acceso a la Universidad.

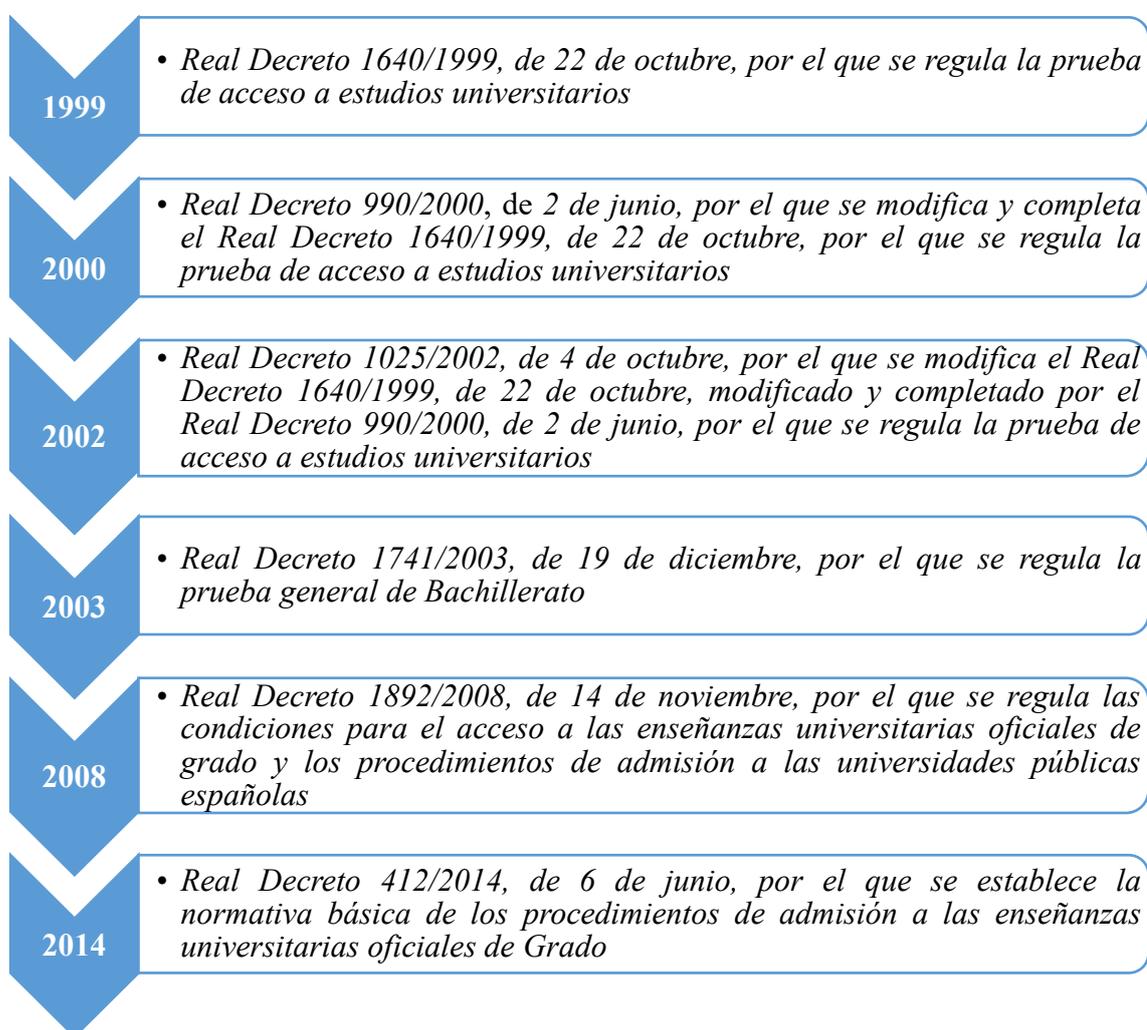


Figura 1.1. Reales Decretos y modificaciones establecidas desde 1999

A lo largo de esta sección centraremos nuestro estudio en realizar una revisión en las Pruebas de Acceso a la Universidad regidas por el Real Decreto 1982/2008, indicando

sus características a nivel normativo y estructural. Para finalizar, y dado que no se cuenta con mucha información fiable, se incluyen algunas pinceladas sobre el nuevo proyecto de las pruebas.

### **1.3.1. Las actuales pruebas 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, el artículo 38 de la Ley Orgánica de Educación (LOE), pone de manifiesto que las pruebas PAU consiste en una prueba de madurez que permite evaluar los conocimientos y capacidades adquiridas por los futuros universitarios durante su etapa de Bachillerato (primer y segundo curso). Con la implantación de este modelo, la selectividad sufrió principalmente cambios en su estructura y puntuación, siendo posible alcanzar una calificación de 14 puntos. Sin embargo, la introducción de estas modificaciones estableció una relación más estrecha entre la citada prueba y los estudios universitarios que se pretenden cursar.

Con los cambios introducidos en la prueba se destacó la desaparición de las vías de examen, siendo los propios estudiantes los que rediseñan la prueba, es decir, los alumnos eligen, las hayan cursado o no, las materias de la modalidad que presentan mayor afinidad con el estudios que desean cursar. El modelo descrito está determinado por dos fases, una de ellas obligatoria para todos los estudiantes que realizan la prueba, denominada por Fase General, y otra, de carácter optativo, Fase Específica, con la que se puede incrementar la calificación final de la prueba PAU hasta 4 puntos. A continuación se detalla cada fase:

*Fase General.* El objetivo de esta fase consiste en la valoración de la madurez y destrezas básicas del estudiante. Consta de cuatro ejercicios que se corresponden a las materias comunes de segundo de Bachillerato. Todos los alumnos se examinarán de una primera prueba relacionada con los contenidos de Lengua Castellana y Literatura; una segunda prueba de Lengua Extranjera y una tercera prueba donde se elegirá entre Historia o Filosofía. Por último, de entre todas las materias ofertadas en la prueba PAU, el alumno seleccionará una de ellas. La calificación de la Fase General es la nota media de los cuatro ejercicios por lo que cada una de las pruebas realizadas supone el 25% de la calificación total. La

calificación máxima que se puede obtener en la Fase General será de 10 puntos. La superación de esta fase permite que el estudiante pueda acceder a los estudios universitarios y la calificación obtenida en esta fase tiene una validez indefinida.

*Fase Específica.* La finalidad de esta fase es la evaluación de los conocimientos en unos ámbitos disciplinares concretos relacionados con los estudios vinculados a la rama de conocimiento que se quiere cursar. Por lo que, la realización de los ejercicios de esta fase permite obtener la Nota de Admisión para cada uno de los títulos de grado a los que el estudiante desea acceder. En esta fase el estudiante seleccionará un máximo de cuatro ejercicios a realizar, considerándose únicamente las dos mejores calificaciones que, una vez ponderadas con un 10% (o 20%), estén relacionadas con la rama del conocimiento de la titulación a la que desea acceder el alumno. A diferencia de la Fase General, las calificaciones de las materias cursadas en esta fase tendrán únicamente una validez de dos cursos académicos siguientes a la superación de la misma.

La nota de acceso a la universidad se obtiene como la media ponderada de la calificación de la Fase General multiplicada por un factor de ponderación del 40% y la calificación media del expediente multiplicada por un factor de ponderación del 60%, es decir,

$$\text{Nota Acceso} = 40\% \text{ Calificación F.G.} + 60\% \text{ Calificación Media Expediente.}$$

La superación de la prueba PAU está condicionada a obtener en la Fase General una calificación igual o superior a 4 y a su vez, la calificación final de la nota de acceso deberá ser mayor o igual a 5 puntos.

Por otro lado, los estudiantes que hayan realizado la Fase Específica podrán optar a una calificación de hasta 14 puntos. En este caso, la calificación nombrada como Nota de Admisión se obtiene como:

$$\text{Nota de Admisión} = \text{Nota de Acceso} + a \times M_1 + b \times M_2,$$

donde  $M_1$  y  $M_2$ , son, de entre todas las materias examinadas en la Fase Específica, las mejores notas de admisión, una vez multiplicadas por los parámetros  $a$  y  $b$ , para cada

una de las carreras solicitadas por el alumno. No obstante, teniendo en cuenta los criterios marcados por las distintas universidades, los factores de ponderación pueden incrementarse un 10%, es decir, si las materias cursadas están estrechamente relacionadas con los futuros estudios universitarios, entonces las universidades tendrán la potestad de considerar una ponderación del 20%. El incremento de los factores de ponderación está condicionado a la existencia de una estrecha afinidad entre la optativa seleccionada y la rama de conocimiento a la que está adscrito el título de grado que se desea realizar.

### **1.3.2. El proyecto en marcha sobre la “nueva selectividad”**

La introducción de la Ley Orgánica para la Mejora de la Calidad Educativa (LOMCE) ha supuesto, entre otros elementos, la eliminación de las Pruebas de Acceso a la Universidad y presenta un nuevo sistema de ingreso a la universidad donde las universidades son las encargadas de realizar la selección de sus futuros estudiantes.

Aunque la LOMCE elimina las PAU, a diferencia de las anteriores leyes promulgadas, introduce una prueba o reválida que es necesaria superar con el fin de obtener el título correspondiente al ciclo de Bachillerato y continuar los estudios universitarios. En el proyecto presentado (MECD, 2015a) se presentan las características generales de la prueba de la evaluación final de Bachillerato y se pone de manifiesto que:

La finalidad de la evaluación final de Bachillerato es comprobar el logro de los objetivos de la etapa y el grado de adquisición de las competencias correspondientes por los alumnos y alumnas en relación con las siguientes materias: a) Todas las materias generales cursadas en el bloque de asignaturas troncales. [...]. b) Dos materias de opción cursadas en el bloque de asignaturas troncales, en cualquiera de los cursos. [...]. c) Una materia del bloque de asignaturas específicas cursada en cualquiera de los cursos, que no sea Educación Física ni Religión (p. 9).

La obtención del título de Bachillerato estará condicionada a la superación de la evaluación final de Bachillerato (obtención de una calificación igual o superior a 5 puntos sobre 10). Por otro lado, la calificación final de esta etapa se deducirá de la siguiente ponderación: a) con un peso del 60%, la media de las calificaciones numéricas obtenidas en cada una de las materias cursadas en Bachillerato; b) con un peso del 40%, la nota obtenida en la evaluación final de Bachillerato.

Con respecto al formato de la prueba, según el primer borrador, la prueba versaría sobre 350 preguntas tipo test de las cuales 200 de ellas estarían relacionadas con materias troncales, 100 con materias optativas y 50 de libre elección por parte del alumno. Sin embargo, esta inicial idea no ha llegado a buen puerto ya que debido a la presión educativa el Ministerio de Educación se ha visto obligado a dar marcha atrás y proponer otro modelo.

Según lo conocido hoy en día, en líneas generales la futura prueba se asemeja a las actuales PAU. La nueva propuesta consta de dos fases, una general que versará sobre las materias obligatorias (de carácter obligatorio), y otra fase específica (de carácter optativo) en la que el estudiante podrá realizar a lo sumo cuatro exámenes seleccionándose únicamente las dos mejores notas. Las preguntas incluidas en cada una de las pruebas tendrán asociadas respuestas a desarrollar. En cuanto a la nota final tampoco se evidencian muchos cambios ya que la fase general seguirá puntuando sobre 10 y los alumnos que realicen la fase específica entonces podrán incrementar su nota hasta 4 puntos.

Entre las novedades recogidas en el nuevo proyecto destacamos el hecho de que las universidades tienen la potestad para llevar a cabo sus propias pruebas de acceso con el fin de seleccionar a sus futuros alumnos. Sin embargo, por el momento, esto son solo conjeturas ya que la mayoría de las universidades descartan la realización de las mismas.

Por último, señalar que aún el Ministerio de Educación, Cultura y Deporte no ha sido publicado las características definitivas de las que consta la mencionada prueba, por lo que seguramente encontraremos cambios nuevos. De lo único que hoy en día se está seguro es del carácter de reválida que tiene.

#### **1.4. IMPORTANCIA DE LA INFERENCIA ESTADÍSTICA**

En la actualidad es habitual encontrar en los medios de comunicación (prensa, radio, TV, libros, Internet) noticias obtenidas a partir de estudios estadísticos. Sin embargo, aunque el ciudadano tiene acceso a toda esa información, para un gran porcentaje de ellos no es nada transparente y en algunas ocasiones difíciles de entender. Los acontecimientos que suceden en la vida diaria ponen de manifiesto la necesidad de

formar a ciudadanos con base estadística para entender correctamente la información recibida. Garfield y Ben-Zvi (2008) señalan que:

Hacer inferencias a partir de los datos es ahora parte de la vida cotidiana y la revisión crítica de los resultados de inferencia estadística a partir de estudios de investigación es una capacidad importante para todos los adultos, (p. 262).

Desde hace años el estudio de la estadística y la probabilidad está recogido en los planes de estudios de todas las etapas educativas, pero sin darle la importancia que tiene en realidad. Sin embargo, es necesario destacar que esta situación está sufriendo grandes cambios. Ejemplo de ello se encuentra en el marco de evaluación de PISA (OCDE, 2015), donde se reconoce la necesidad de que los estudiantes desarrollen las capacidades necesarias para enfrentarse, en su vida diaria y profesional, a la toma de decisiones, tanto desde una perspectiva matemática como científica, en situaciones en las que se carece de información.

La inferencia estadística es considerada una rama de la Estadística que permite obtener conclusiones acerca de una población, a partir de la información proporcionada por una muestra representativa de la misma. Desde el punto de vista formal, la inferencia consiste en un conjunto de métodos que permiten, estimando sus parámetros con un grado de confianza o contrastando hipótesis sobre los mismos, generalizar los resultados obtenidos en una muestra a una población.

A continuación analizaremos la importancia que juega la inferencia estadística en el desarrollo de otras ciencias y en la formación del estudiante.

#### **1.4.1. Papel en el desarrollo de las ciencias**

Respecto al origen de la inferencia, Hald (2008) indica que los primeros trabajos se deben a Bayes y Laplace que, independientemente, resuelven entre 1764 y 1786 el llamado problema de la probabilidad inversa, germen de la inferencia bayesiana. Entre 1809 y 1828 Laplace y Gauss encuentran la distribución muestral de las medias (que se aproxima mediante la distribución normal).

Pero el grueso de la teoría de la inferencia estadística se creó como respuesta al problema consistente en obtener un conocimiento general (una nueva teoría científica) a partir del análisis de casos particulares (inducción empírica). Esto es, nace con la búsqueda de métodos que permitan justificar el razonamiento inductivo y la extensión

de sus conclusiones, problema de gran importancia en las ciencias empíricas. Este problema ha ocupado a los filósofos y estadísticos por largo tiempo, sin que se haya obtenido una solución aceptada por consenso (Batanero y Díaz, 2006; Cabriá, 1994; Rivadulla, 1991).

Popper, en su libro *The logic of scientific discovery*, escrito en 1934 y vuelto a reescribir en 1959 propuso aceptar una teoría como provisionalmente cierta si, a pesar de numerosos intentos, no se consigue refutarla. Este autor sugirió poner a prueba las hipótesis científicas, mediante experimentos u observaciones y comparar los patrones deducidos de la teoría con los datos obtenidos. La teoría sería provisionalmente confirmada si los datos recogidos siguiesen estos patrones. Estas ideas de Popper tuvieron una gran influencia en el desarrollo de la inferencia estadística. Puesto que mediante un razonamiento inductivo no es posible llegar a la certidumbre de una proposición (verdad cierta), diversos autores intentaron calcular la probabilidad de que una hipótesis sea cierta (verdad probable) (Batanero, 2000; Rivadulla, 1991).

Las soluciones alcanzadas llevan a diferentes metodologías de la inferencia que coexisten en la actualidad. De ellas, consideramos en este trabajo la frecuencial, que es la considerada en las PAU; aunque en otros países se incluye en Bachillerato la bayesiana y la basada en simulación intensiva (o re-muestreo), todas desde un enfoque informal.

En la primera, que es la que actualmente se considera en Bachillerato en España, se supone que un parámetro es un valor fijo que describe una distribución de probabilidad en una población y se estima el valor del parámetro a partir de datos de una muestra. Respecto a los contrastes estadísticos, encontramos dos aproximaciones: (a) las pruebas de significación, que fueron introducidas por Fisher y (b) los contrastes como reglas de decisión entre dos hipótesis, que fue la concepción de Neyman y Pearson. El test de significación fue propuesto por Fisher (1935) en su libro *The design of experiments*, publicado en 1935, y es un procedimiento que permite rechazar una hipótesis, con un cierto nivel de significación. Neyman y Pearson (1933) por su parte estaban interesados en el contraste de hipótesis como un proceso de decisión que permite elegir entre una hipótesis dada  $H_0$  y otra hipótesis alternativa  $H_1$  (Rivadulla, 1991). Además, los autores introducen la idea de error tipo II (probabilidad de no rechazar la hipótesis nula sabiendo que es falsa), no tenida en cuenta por Fisher.

El hecho de poder realizar generalizaciones sobre una población, ha provocado que durante las últimas décadas la inferencia estadística adquiriera una gran importancia generalizándose su enseñanza en la mayoría de los grados y másteres universitarios en las universidades. A modo de resumen se presenta en la Figura 1.2 un esquema en el que se refleja la Inferencia Estadística en su enfoque frecuencial y en relación con la probabilidad.

Por otro lado, en la metodología Bayesiana (Lecoutre y Poitevineau, 2014), el parámetro es una variable aleatoria con distribución inicial conocida y la finalidad de la inferencia es usar el teorema de Bayes y los datos recogidos de una muestra para llegar a la distribución final (o actualizada) del parámetro. En los métodos de re-muestreo no se considera una distribución teórica en la población. Se parte de una muestra de datos y se utiliza esta muestra para simular un gran número de muestras (obtenidas de la primera). La finalidad es estimar la variabilidad de un estadístico en el muestreo y a partir de ello inferir el valor del parámetro en la población (Batanero, 2015).



Figura 1.2. Esquema de Inferencia Estadística

Fuente: Elaboración propia

### **1.4.2. Importancia en la formación del estudiante**

En la formación de los alumnos, la importancia del estudio de la inferencia queda reflejada en los contenidos incluidos en la normativa curricular vigente de Bachillerato y la influencia que su conocimiento tienen en la nota conseguida en las pruebas PAU. Las normativas curriculares señalan que una buena formación en inferencia estadística capacita a cualquier ciudadano para:

- analizar, interpretar y valorar fenómenos sociales, con objeto de comprender los retos que plantea la sociedad actual;
- formar criterios propios sobre fenómenos sociales y económicos;
- interpretar de datos y mensajes, argumentando con precisión y rigor;
- formular de hipótesis que permiten contrastar estrategias para la resolución de problemas;
- justificar de procedimientos aportando rigor a los razonamientos, detectando inconsistencias lógicas;
- adquirir y manejar con fluidez un vocabulario específico de términos y notaciones matemáticos;
- establecer relaciones entre las matemáticas y el entorno social, cultural o económico.

Una adecuada formación en inferencia estadística contribuye a formar a ciudadanos autónomos que pueden dar respuesta a muchas de las necesidades que requiere la sociedad en la que vivimos actualmente. Es por ello que el documento *Guidelines for Assessment and Instruction in Statistics Education, GAISE* (Franklin et al., 2007), de la American Statistical Association y el National Council of Teachers of Mathematics recomienda que los estudiantes aprendan a mirar detrás de los datos e interpretar la variabilidad, reconozcan la posibilidad de realizar inferencias a partir de nuestras aleatorias o de experimentos bien diseñados.

Tukey (1962, p. 13) señala que “Mucho mejor una respuesta aproximada a una pregunta correcta, que es a menudo vaga, que una respuesta exacta a la pregunta errónea, que puede hacerse siempre de forma precisa”.

## 1.5. OBJETIVOS E HIPÓTESIS DE LA INVESTIGACIÓN

Finalizada la exposición del tema y la justificación de su importancia, pasamos a describir los objetivos que se persiguen con la presentación de la memoria y se formulan algunas hipótesis sobre lo que esperamos encontrar.

### Objetivos del trabajo

Tal y como ha sido señalado en la introducción, el objetivo principal del trabajo es la realización de un estudio sobre los problemas de inferencia estadística que han sido propuestos en las Pruebas de Acceso a la Universidad (PAU) durante el periodo 2003-2014 en la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II. El objetivo principal del trabajo puede desglosarse en los siguientes objetivos específicos:

OE1. *Revisar las dificultades en la comprensión de los conceptos asociados a la inferencia estadística.* Previo al estudio empírico, es importante analizar las investigaciones previas, para fijar nuestro trabajo en un contexto más amplio. El interés de este objetivo radica en que el conocimiento sobre los errores más frecuentes o dificultades que tienen los estudiantes cuando trabajan los contenidos de inferencia estadística, supone una ayuda a los docentes. El hecho de contar con esta información, permite a los profesores adaptar su proceso de enseñanza-aprendizaje con el fin de que los estudiantes asimilen, de una forma correcta, los contenidos incluidos en el bloque objeto de estudio. Para lograr este objetivo se realiza una síntesis de investigaciones previas en el Capítulo 2.

OE2. *Analizar el contenido sobre inferencia estadística propuesto en PAU.* Los problemas propuestos a los estudiantes, bien durante la enseñanza, o en las pruebas de evaluación, van a incidir directamente en el aprendizaje de diferentes objetos matemáticos y el significado asignado a los mismos. El enfoque Onto-Semiótico (Godino, Batanero y Font, 2007), asume que los objetos matemáticos surgen de las prácticas matemáticas realizadas para resolver problemas relacionados con el objeto, considerando el término problema en sentido amplio. Por ello, es importante conocer con el mayor detalle posible las características de la evaluación. Para lograr este objetivo y en base al enfoque Onto-Semiótico, se han resuelto cada uno de los problemas las pruebas PAU para Matemáticas Aplicadas a las Ciencias Sociales II

durante el periodo 2003-2014 con el fin de analizar los objetos relacionados con la inferencia estadística y determinar el significado evaluado en las pruebas.

OE3. *Identificar las variables que definen los problemas de inferencia estadística y analizarlas estadísticamente.* La investigación sobre resolución de problemas muestra que algunas variables de los mismos inciden directamente en su dificultad. Es por ello importante identificar las principales variables que aparecen en los problemas propuestos a los estudiantes sobre inferencia estadística. Para lograr este objetivo, partiendo de la información ofrecida por el objetivo específico anterior, OE2, se identifican distintos objetos relacionados con la inferencia estadística los cuales permitirán describir las variables que serán objeto de estudio. Posteriormente, se codifican los datos para presentar la distribución de cada una de las variables globalmente y por año.

### **Hipótesis o expectativas iniciales**

Para finalizar, se formulan algunas hipótesis, que se entiende cómo expectativas iniciales y que se discutirán en las conclusiones. Señalar que no realizaremos contrastes de hipótesis en el sentido formal de la inferencia.

H1: *El análisis de las investigaciones previas mostrará que el tema que nos ocupa ha sido escasamente investigado y, por tanto, nuestro trabajo puede aportar información original.* Para formular esta hipótesis nos basamos en nuestra primera búsqueda bibliográfica, en la que sólo tres trabajos están ligeramente relacionados con el actual. Igualmente nos apoyamos en el estudio de los antecedentes realizado por Carretero (2014), que tampoco encontró investigaciones similares a la que ahora abordamos.

H2: *El análisis semiótico de los problemas de inferencia propuestos en las PAU y su comparación con el contenido curricular de estadística en los decretos curriculares de Bachillerato mostrará, por un lado, que se da una alta importancia a la inferencia en las PAU. Además se observará un buen ajuste entre los contenidos de inferencia fijados en el currículo y los considerados en las pruebas.* Nos basamos para formular esta hipótesis en un análisis preliminar publicado por Díaz, Mier, Alonso y Rodríguez-Muñiz (2014).

*H3. Aunque en el currículo que han cursado los estudiantes en los pasados años (MEC, 2007) se sugiere estudiar la inferencia tanto para una como para dos muestras y en relación a la media y proporción, pensamos que la mayoría de problemas se centrarán tan sólo en la inferencia sobre la media de la muestra. Formulamos esta hipótesis, debido que estos problemas son más sencillos, en comparación con los otros citados.*

*H4. Al estudiar el contexto de los problemas, pensamos que se encontrarán una proporción grande de problemas descontextualizados. Esta hipótesis se confirmó en el estudio que realiza Carretero (2014) sobre los problemas de probabilidad y pensamos se repetirá en nuestro estudio.*

# Capítulo 2

## FUNDAMENTOS

### Índice de Contenidos

---

- 2.1. Introducción
  - 2.2. Marco Teórico
  - 2.3. Investigaciones sobre las Pruebas de Acceso a la Universidad
  - 2.4. Análisis sobre otras pruebas de evaluación
  - 2.5. Dificultades de los estudiantes con la inferencia
  - 2.6. Conclusiones
- 

### 2.1. INTRODUCCIÓN

Una vez mostrados los objetivos que se persiguen con el desarrollo de esta memoria y las hipótesis iniciales sobre los resultados que esperamos obtener en la investigación, se presenta en este capítulo los fundamentos que sustentan el trabajo.

Puesto que el marco teórico que ha sido empleado es el Onto-Semiótico, se resumen algunos elementos de dicho enfoque. Considerando los objetivos específicos descritos en el Capítulo 1, hemos centrado nuestra atención en el análisis del significado institucional, ya que nos permitirá analizar la comprensión del tema de inferencia estadística por parte del estudiante.

En el segundo apartado del Capítulo, se realiza una revisión de las investigaciones más relacionadas con nuestro trabajo que se han llevado a cabo con el fin de analizar las Pruebas de Acceso a la Universidad. Puesto que el tema que estamos tratando forma

parte del área de Matemáticas, presentamos, por un lado, los antecedentes de investigaciones relacionadas con análisis de las pruebas de matemáticas y se prosigue con aquellos trabajos que han centrado su atención en el estudio de la inferencia estadística en las PAU. Seguidamente, y en la misma línea que el punto anterior, se realiza un barrido de algunas investigaciones que han tenido como objetivo el análisis de otras pruebas de evaluación distintas a las PAU.

Dentro del proceso de enseñanza y aprendizaje de las matemáticas un elemento a tener en cuenta, a la hora de tratar cualquier tema, son las posibles dificultades que pueden tener los estudiantes. El estudio o análisis de dichas dificultades puede contribuir a la mejora de la enseñanza y aprendizaje. En nuestro caso, aunque el objetivo es el análisis de los problemas de inferencia planteados en las PAU de Andalucía, en la valoración de la complejidad los mismos es importante tener en cuenta las posibles dificultades de los estudiantes al resolverlos. Por ello, se muestran algunas investigaciones relacionadas con las dificultades de los estudiantes en el tema en cuestión.

## **2.2. MARCO TEÓRICO**

Nos apoyamos en herramientas teóricas propuestas por el Enfoque Onto-Semiótico (EOS) del conocimiento y la instrucción matemáticos, sistema teórico elaborado Godino, Batanero y Font (2007). Como formación epistemológica, dicho enfoque asume que las matemáticas provienen de la actividad humana orientada a la resolución de determinados tipos de problemas, los cuales constituyen la razón de ser y significado de los objetos emergentes de la misma.

En el EOS se 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, considerando el término problema en sentido amplio (incluyendo también ejercicios y tareas). El sentido dado a este término por los autores es muy amplio y puede englobar cualquier tarea o ejercicio que requiera el uso de las matemáticas en su resolución. Ejemplos de problemas relacionados con la inferencia y propuestos en las PAU 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 (Batanero y Borovcnick, 2016).

Seguidamente, se describen los significados institucionales y personales, señalando los distintos tipos; cómo el EOS considera la actividad de evaluación y para concluir, se presenta la idoneidad didáctica y sus componentes.

### **2.2.1. Significados institucionales y personales y sus tipos**

Mediante la reiteración de las prácticas matemáticas realizadas al resolver muchos problemas similares, se llega a configurar el significado del objeto matemático o conjunto de prácticas asociado al mismo. Por tanto los problemas propuestos a los estudiantes, bien durante la enseñanza o en las pruebas de evaluación van a incidir directamente en su adquisición de diferentes objetos matemáticos y el significado asignado a dichos objetos y de ello se deduce la importancia del análisis de tales problemas.

Dentro del marco teórico EOS se diferencia 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) (véase Figura 2.1).

- Dentro de la institución de enseñanza, se pueden distinguir cuatro tipos de significados: referencial, pretendido, implementado y evaluado. El *significado referencial* sería el más amplio y, en nuestro caso, sería el correspondiente a la inferencia en la propia estadística. Atendiendo a diversos factores de las orientaciones curriculares (por ejemplo, el nivel educativo, la temporalización, conocimientos previos de los estudiantes, etc.) se concreta el *significado pretendido*. El significado pretendido del objeto dentro de un nivel educativo sería el fijado por la institución. El *significado implementado* es el que el profesor presenta a los estudiantes en cada centro o aula y el *significado evaluado* sería el contenido en las pruebas de evaluación.
- De igual modo se distinguen tres tipos de significado personal de un objeto matemático. Se identifica el *significado global* que corresponde con todo el sistema potencial de prácticas subjetivas; el *significado declarado* asociado a las

que se pueden identificar al realizar una evaluación y el *significado logrado* relacionado con las prácticas que coinciden con las pretendidas institucionalmente (Font, Godino, y Gallardo, 2013).

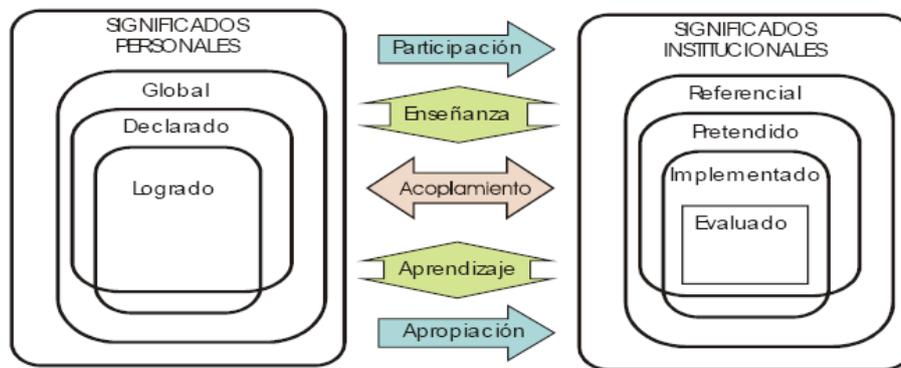


Figura 2.1. Tipos de significados institucionales y personales (Godino, 2002, p.141).

En nuestro caso, el estudio se orientará a determinar el significado institucional de la inferencia estadística evaluado en las Pruebas de Acceso de Matemáticas Aplicadas a las Ciencias Sociales II. La importancia de este punto se fundamenta en que las respuestas a las pruebas de evaluación nos 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 permitirnos adquirir 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 trataremos de evaluar la adecuación de esa correspondencia para el contenido pretendido de la inferencia en las orientaciones curriculares de Bachillerato.

### 2.2.2. Evaluación en el EOS

Al considerar la actividad de evaluación, Godino (1996) sugiere que es necesario considerar dos ejes:

- Eje descriptivo, que indica los aspectos o componentes de los objetos que se pretende evaluar, en nuestro caso, los contenidos de inferencia que se incluyen en las PAU.
- Eje procesual, que indica las fases o niveles necesarios para considerar que un estudiante alcanza la comprensión; en las PAU será el tipo de respuestas consideradas adecuadas en cada tipo de problema.

Tal y como se mencionó anteriormente, el problema de la evaluación de la comprensión personal sobre cierto objeto matemático, como la inferencia, se debe a que la comprensión no se puede observar directamente. Este tipo de evaluación se debe deducir de las respuestas en las pruebas de evaluación que tiene el estudiante, es decir, de sus prácticas personales.

Dentro del Enfoque Onto-Semiótico, la evaluación de la comprensión sería el estudio de la correspondencia entre los significados personales e institucionales. Por tanto, la comprensión depende de la institución desde donde se lleva a cabo la evaluación, que es la que determina hasta qué punto un sujeto “comprende” el significado de un objeto. En nuestro caso son los responsables de establecer las pruebas los que deciden los criterios para asegurar la comprensión de la inferencia por parte de los estudiantes.

### **2.2.3. Idoneidad didáctica y sus componentes**

Otro elemento del marco teórico que usaremos para valorar nuestras conclusiones es el de idoneidad didáctica, que Godino, Contreras y Font (2006) definen como un indicador de la calidad de un proceso educativo (o bien un material o un instrumento de evaluación, en nuestro caso), que se compone de seis componentes (Godino, 2011):

- *Idoneidad epistémica*: Indica el grado de concordancia entre los significados institucionales implementados (pretendidos o evaluados), respecto de un significado de referencia. En nuestro caso, la concordancia entre el significado de la inferencia pretendido en el currículo y el evaluado en las pruebas PAU.
- *Idoneidad cognitiva*: grado en que los significados pretendidos/ implementados son asequibles a los estudiantes. En el estudio que estamos realizando, no los podremos deducir directamente, sino por la información de otros trabajos realizados con

alumnos, o indirectamente, mediante el análisis de la dificultad potencial de los problemas.

- *Idoneidad interaccional*: grado en que las configuraciones y trayectorias didácticas permiten identificar conflictos semióticos potenciales y resolver conflictos observados durante el proceso de instrucción. Este componente no se aplica en nuestro estudio.
- *Idoneidad mediacional*: grado de disponibilidad y adecuación de los recursos materiales y temporales necesarios para el desarrollo del proceso de enseñanza-aprendizaje. En estas pruebas el principal material sería el tiempo disponible, así como las calculadoras y las tablas de probabilidades de la distribución normal que se les proporcionan al alumno.
- *Idoneidad emocional*: grado de implicación (por ejemplo interés o motivación) del alumnado en el proceso de estudio. Lo valoraremos analizando los contextos de los problemas y su posible interés para el estudiante.
- *Idoneidad ecológica*: grado en que el proceso de estudio se ajusta al proyecto educativo de la institución (o comunidad) y a los condicionamientos del entorno en que se desarrolla (Godino, Bencomo, Font y Wilhelmi, 2006).

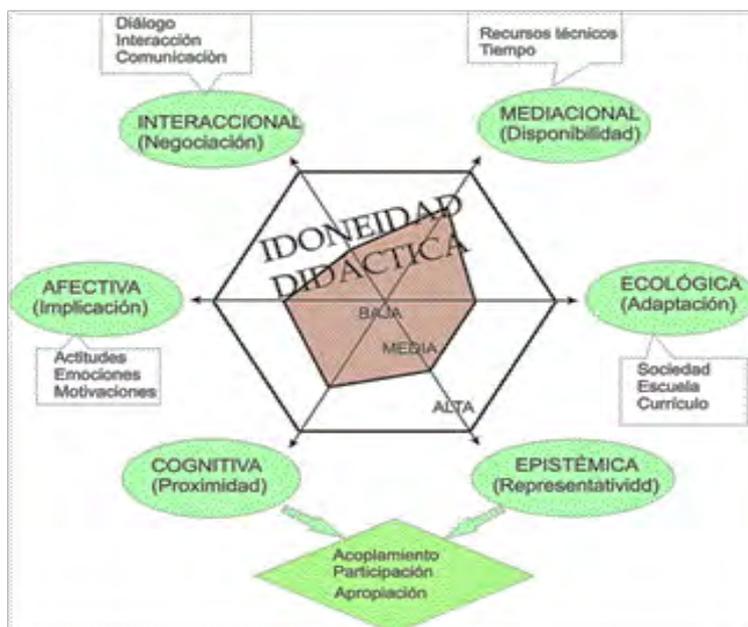


Figura 2.2. Idoneidad Didáctica (Godino, 2011, p. 6)

## **2.3. INVESTIGACIONES SOBRE LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD**

Desde la implantación de las Pruebas de Acceso a la Universidad son varios los trabajos que han abordado, desde diferentes puntos de vista, el estudio de las mismas. En el trabajo presentado por Ruiz, Dávila, Etxeberría y Sarasua (2013), los autores realizan una clasificación de las distintas investigaciones que se han llevado a cabo sobre dichas pruebas. A grandes rasgos se distinguen las siguientes líneas de investigación:

- estudio de los resultados obtenidos por los estudiantes en dichas pruebas en algunas universidades, analizando las calificaciones en diferentes materias o comparando universidades entre sí;
- comparación de las pruebas con otros métodos para evaluar a los alumnos en su entrada a la universidad en otros países;
- estudios sobre los resultados y aspectos didácticos de alguna materia en particular de la prueba;
- análisis de algunos de los factores que inciden en el rendimiento y la evaluación de los alumnos en las pruebas de acceso, así como los aspectos relativos a la validez y fiabilidad de las pruebas.

A lo largo de este apartado se realizará una revisión bibliográfica de las principales investigaciones realizadas en esta línea. Puesto que el Trabajo Fin de Máster pertenece al área de Didáctica de la Matemática, se ha visto conveniente analizar los estudios llevados a cabo sobre las pruebas de matemáticas para posteriormente centrarnos en aquellos que tratan el tema de la inferencia estadística en las mencionadas pruebas.

### **2.3.1. Análisis de las pruebas de matemáticas**

En la revisión sobre las investigaciones centradas en la asignatura de Matemáticas I y/o Matemáticas Aplicadas a las Ciencias Sociales II, se ha podido constatar la gran variedad de trabajos realizados. Con respecto a estas materias generalmente se han analizado los resultados de los estudiantes (ej., Blázquez y Luengo, 1989), la opinión de los profesores (Murillo, 1997; Ruiz, Dávila, Etxeberría y Sarasua, 2013), las características psicométricas de las prueba (ej., Gaviria, 2005), la evaluación de la

calidad de la corrección de las PAU (Cuxart, 2000; Grau, Cuxart, y Martí-Recober, 2002; Mengual, Gorgorió y Albarracín, 2013; Nortes, Nortes y Lozano, 2015) o las competencias matemáticas evaluadas (Boal, Bueno, Lerís y Sein-Echaluce, 2008).

Murillo (1997) entre los diversos estudios que realiza, muestra un análisis de las pruebas correspondientes a la convocatoria de junio durante el curso académico 1994-1995 de los 29 distritos universitarios. El autor señala que la asignatura Matemáticas II (actualmente denominada Matemáticas Aplicadas a las Ciencias Sociales II) es la asignatura que presenta mayor dificultad a los estudiantes desde el punto de vista del porcentaje de aprobados ya que más de la mitad de los estudiantes no llegaron a superar el examen. Por otro lado, en el trabajo se pone de manifiesto que la mayoría de las preguntas incluidas en las pruebas tienen un carácter práctico, y que los contenidos abordados son los asociados al Cálculo de probabilidades y Sistemas de ecuaciones lineales.

En 2010, Contreras, Ordóñez y Wilhelmi centra su investigación en los significados de la integral definida en una muestra de libros de texto y en 140 pruebas de selectividad propuestas en la comunidad andaluza desde 1999 hasta 2008. A partir de dicho estudio, los autores obtienen una descripción de la realidad institucional con relación a la integral definida, que permite sugerir acciones para la mejora de los procesos de enseñanza y aprendizaje relacionado con el cálculo integral que se aborda en el segundo curso de Bachillerato de la asignatura Matemáticas I. Entre los resultados obtenidos, los autores destacan una reducción generalizada de la evaluación en las PAU a métodos algebraicos y un aislamiento de las técnicas de resolución, indicando pues una baja idoneidad epistémica.

Nortes y Nortes (2010), analizan las respuestas de una muestra de alumnos a la materia Matemáticas Aplicadas a las Ciencias Sociales de la PAU celebradas en la Universidad de Murcia, en septiembre de 2009. En el trabajo, los autores informan sobre las calificaciones; también detectan los errores más significativos y comparan los estadísticos de cada una de las cuestiones y de los bloques.

Ruiz, Dávila, Etxeberria y Sarasua (2013) realizan un análisis estadístico de los resultados individuales obtenidos en Matemáticas por los alumnos en el País Vasco en el curso 2009-2010, comparando estos resultados por centro y analizando la relación de estos resultados con la opinión de los profesores sobre los mismos y la metodología del

proceso de enseñanza-aprendizaje utilizada. Encontraron unos resultados muy pobres con pocas diferencias entre centros educativos. Un resultado no esperado es que los centros con buenos resultados utilizaban una metodología tradicional, con poca presencia de las nuevas tecnologías.

Mallart (2014) analiza las respuestas de 104 estudiantes en la asignatura de Matemáticas de las Pruebas de Acceso a la Universidad de 2012, estudiando las resoluciones con el propósito de detectar errores significativos y poder incidir en tales dificultades de aprendizaje. Diferencia entre resoluciones mecánicas, la ejecución con precisión de las operaciones, la aplicación con rigor de las propiedades y la capacidad creativa de resolución. El estudio realizado muestra que las primeras son más frecuentes, lo que indica que los alumnos han aprendido los aspectos rutinarios en la resolución de problemas, pero se observa falta de creatividad, de rigor y precisión en los problemas geométricos.

Para finalizar, destacar que aunque se han realizado diversas investigaciones sobre las PAU, pocas han estado centradas en el análisis de los problemas planteados. Después de una profunda revisión, se han encontrado como trabajos más afines a la investigación que se está llevando a cabo los presentados por Ruiz, Sarasua y García (2011) y Díaz, Mier, Alonso y Rodríguez-Muñiz (2014).

Ruiz y colaboradores (2011) analizan los ejercicios que han sido propuestos en las PAU desde 1994 hasta 2008 en la Universidad del País Vasco. El estudio tiene como objetivo establecer una clasificación sobre los problemas-tipo análogos en cuanto a su método de resolución de las materias de Matemáticas II y Matemáticas Aplicadas a las Ciencias Sociales II. Concretamente, los autores realizan una clasificación según el nivel de dificultad (estándar y difícil) y según el contenido tratado en el mismo. Los autores comprueban que, según la clasificación realizada, un alto porcentaje de problemas de tipología difícil estaban relacionados con una calificación baja y un número alto de problemas de tipología estándar estaban asociados a calificaciones buenas. De los resultados, los autores recomiendan que si se incluyeran más problemas estándar en las Matemáticas de Ciencias Sociales II, seguramente mejorarían los malos resultados que tienen los estudiantes en esta prueba.

En la misma línea, Díaz y colaboradores (2014) realizan un estudio comparativo de cómo las pruebas PAU de cuatro Comunidades Autónomas tratan los contenidos de

estadística y probabilidad, concluyendo que la distribución de las unidades curriculares en los exámenes es muy heterogénea.

Por parte del Grupo de Investigación sobre Educación Estadística de la Universidad de Granada, se han realizado diversos trabajos enfocados en el análisis de los problemas de probabilidad propuestos en las Pruebas de Acceso a la Universidad en Andalucía en el periodo 2003-2014 (ej., Carretero, López-Martín y Arteaga, 2015; Contreras, López-Martín, Arteaga y Carretero, 2015; López-Martín, Contreras, Batanero y Carretero, 2015; López-Martín, Contreras, Carretero y Serrano, 2016). Los resultados obtenidos en estos trabajos 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 presente memoria tiene como finalidad complementar los trabajos enfocados al análisis de los problemas de probabilidad de las PAU, para ofrecer una visión de la evaluación dedicada al bloque de estadística y probabilidad dentro de las Pruebas PAU en Andalucía.

### **2.3.2. Estudio de la inferencia en las PAU**

Tal y como se ha puesto de manifiesto en el apartado anterior, la búsqueda de investigaciones relacionadas con el presente trabajo ha estado muy limitada. Las investigaciones más cercanas al análisis de los problemas de inferencia estadística en las pruebas PAU son las llevadas a cabo por Espinel, Ramos y Ramos (2006), Espinel, Ramos y Ramos (2007) y García y García (2005).

En el trabajo presentado por García y García (2005), los autores analizan las respuestas dadas por 50 estudiantes a los problemas de inferencia estadística de la convocatoria de junio de 2002 en la Universidad de la Laguna. Concretamente, los autores cruzan la solución del profesorado corrector con la forma de llevar a cabo el planteamiento, los cálculos y la resolución los estudiantes. Los resultados muestran que aproximadamente un 72% de los estudiantes seleccionan los dos ejercicios de inferencia estadística, sin embargo, solo siete alumnos consiguen concluir correctamente el test de hipótesis y determinar el intervalo de confianza.

Por otra parte, Espinel, Ramos y Ramos (2006) seleccionan 272 estudiantes que realizaron la PAU en la Universidad de la Laguna en 2005 con el objeto de analizar la tendencia que tienen los estudiantes de elegir las actividades relacionadas con estadística y probabilidad. De los resultados se observa que las preguntas con contenidos estadísticos son las más seleccionadas por los estudiantes (90%) y corresponden a las calificaciones más elevadas, mientras que los ejercicios de funciones son los menos seleccionados (20%) y van asociados a las calificaciones más bajas.

Empleando la misma muestra, Espinel, Ramos y Ramos (2007) ampliaron el estudio realizando una revisión sobre los errores más comunes cometidos por los estudiantes al abordar cuestiones de inferencia estadística. Comparando los resultados con investigaciones previas, los autores señalan que el método de enseñanza-aprendizaje de la estadística no se ha adaptado lo suficientemente siendo necesario incluir ciertas mejoras. Por tal motivo, los autores proponen un desarrollo metodológico basado en: 1) uso de analogías en el desarrollo de los contenidos teóricos; 2) manejo de las TICs; 3) trabajo con proyectos estadísticos.

Destacamos que en las investigaciones halladas, los autores ponen el foco de atención en la forma de actuar de los estudiantes y en el estudio de las dificultades asociadas a la comprensión de conceptos estadísticos. Señalamos que en estas investigaciones no se lleva a cabo un estudio o clasificación de los problemas de inferencia, ni tratan de identificar las variables que los definen. En base a ello se considera que nuestro trabajo aporta información original.

#### **2.4. ANÁLISIS SOBRE OTRAS PRUEBAS DE EVALUACIÓN**

Entre los investigadores que se han interesado por analizar pruebas de evaluación encontramos a Rico (2006a), 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. En cada uno de los problemas analizó las variables: contenido matemático; contexto o situación; nivel de complejidad y procesos necesarios para realizar una tarea matemática. La autora concluye la investigación señalando el desequilibrio existente en la distribución de los ítems respecto al contexto y al contenido. 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 enfocados a estudiantes de último grado del ciclo de educación básica primaria en los años 2003, 2006 y 2009. La autora analiza el tipo de gráfico empleado o tabla empleada (como por ejemplo el diagrama de sectores o el pictograma); la competencia evaluada (razonamiento, resolución o comunicación); el nivel de lectura de gráficos (tipo de dato que debe obtener del gráfico); 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 otros gráficos recomendados en el currículo.

## **2.5. DIFICULTADES DE LOS ESTUDIANTES CON LA INFERENCIA**

La inferencia es el tema de estadística en que probablemente hay mayor número de investigaciones sobre errores de comprensión. Una enseñanza desde un enfoque formal deductivo y la necesidad de que el estudiante domine nociones tan abstractas como por ejemplo, distribución muestral, intervalo de confianza, estadístico de contraste, hipótesis

nula y alternativa, nivel de confianza y significación, presentan grandes dificultades para los estudiantes (Harradine, Batanero y Rossman, 2011).

Son muy pocos los trabajos realizados con estudiantes de Bachillerato. Destacamos el llevado a cabo por Espinel, Ramos y Ramos (2007) donde las autoras realizan una lista de los errores más frecuentes que cometen los estudiantes cuando abordan las cuestiones de inferencia estadística incluidas en las PAU de La Laguna. Los errores que detectan son: 1) dificultad al entender la noción de probabilidad; 2) tipificación errónea de la distribución Normal; 3) uso incorrecto de la tabla de probabilidades de la Normal; 4) fallo en la aproximación de la Binomial a la Normal; 5) confusión entre muestra y población; 6) descuido ante cambios del tamaño muestral; 7) dificultad para discernir entre parámetro, estimador y estimación; 8) identificación y formulación confusa de los contrastes unilaterales y bilaterales; 9) intercambio del valor empírico con el valor muestral; 10) asignación incorrecta del nivel de confianza; 11) asociación errónea del nivel de significación y la probabilidad buscada; 12) interpretación errónea de las conclusiones en los contrastes. Estos errores son analizados con detalle en Ramos, Espinel y Ramos (2009).

A continuación se describen los errores más comunes en la inferencia estadística. De las investigaciones desarrolladas en este campo, se han seleccionado aquellas que centran su atención en alumnos universitarios. La elección de este tipo de trabajos nos permitirá valorar la posible dificultad de las PAU. Una información más amplia se puede encontrar en Batanero y Borovcnik (2016), Batanero y Díaz (2006), Castro-Sotos, Vanhoof, Noortgate y Onghena (2007), Harradine, Batanero y Rossman (2011) y Vallecillos (1994, 1999).

### **2.5.1. Comprensión de la distribución muestral**

Al iniciar el estudio de la inferencia estadística, los conceptos de población y muestra, junto a sus características, adquieren una gran importancia. El proceso de muestreo, que permite obtener, al azar, un valor de la variable, debe comprenderse como un sistema de ideas interrelacionadas (selección al azar, replicación, variabilidad y distribución), para lo cual es esencial la identificación de la distribución en muestreo. Sin embargo, tal identificación genera ciertas confusiones, ya que es necesario manejar simultáneamente

los conceptos de distribución estadística (datos), distribución de la variable aleatoria (población) y distribución muestral, conceptos que a veces se confunden (Harradine, Batanero y Rossman, 2011).

Inicialmente, se comienza considerando una población con una variable y la *distribución de la variable en la población*; por ejemplo la altura de chicos de 18 años en un cierto país o ciudad. En esta población nos interesa conocer un parámetro, por ejemplo, su media  $\mu$  (altura media de todos los chicos en la población). Para estimarla tomamos una muestra aleatoria de datos; por ejemplo 500 chicos a los que se mide su altura. A partir de estas medidas se obtiene la *distribución de la variable estadística* “altura de los 500 chicos de la muestra” y la media de dicha muestra,  $\bar{x}$ , será empleada para estimar la media de la población. A partir del ejemplo empleado, es evidente que es necesario trabajar con dos medias diferentes; la de la población y la de la muestra. Mientras que la primera es una constante desconocida (no se conoce el verdadero valor de la media en la población) la segunda es un valor conocido en nuestra muestra.

Ahora bien, cuando consideramos no sólo la muestra utilizada, sino todas las posibles muestras de 500 elementos tomadas aleatoriamente de la misma población, al calcular la media de cada una de estas muestras, tenemos una variable aleatoria, porque la media varía de una muestra a otra. La distribución de esta nueva variable aleatoria (media de todas las posibles muestras del tamaño dado) se llama *distribución muestral*. En el caso particular del ejemplo que hemos tratado, la teoría estadística indica que la distribución muestral es una distribución normal, cuya media coincidiría con la media de la variable de partida en la población.

En definitiva, tenemos tres distribuciones diferentes y tres medias diferentes, que habitualmente los alumnos confunden (Schuyten, 1991). El trabajo con estos tres niveles y el uso de cada uno en el proceso de inferencia tiene una gran complejidad semiótica y por tanto, no es de extrañar que los alumnos cometan errores, como se muestra en algunas investigaciones. Además, Liu y Thompson (2009) indican que los estudiantes sólo consideran la muestra dada en los ejercicios que resuelven, sin ser capaces de imaginar las infinitas muestras potenciales que pueden tomarse de una población; por ello no ven a la media de la muestra como una variable aleatoria.

### 2.5.2. Comprensión del contraste de hipótesis

La realización de un contraste de hipótesis está influenciada por el manejo y comprensión de una gran variedad de conceptos y términos que son necesarios para obtener una correcta aplicación e interpretación del mismo (Inzunza y Jiménez, 2013; Vallecillos, 1994). Por ello, es fácil encontrar dentro de la investigación sobre didáctica de la estadística una gran variedad de trabajos en los que se han descrito errores e interpretaciones incorrectas relacionadas, en general, con la comprensión del contraste de hipótesis (Batanero, 2000; Falk y Greenbaum, 1995). Algunas de estas dificultades tienen su origen, en parte, en una deficiente comprensión del concepto de probabilidad condicional, pues algunos conceptos de inferencia se definen usándola (Falk, 1986).

Un concepto que suele tener una interpretación incorrecta es el nivel de significación, denotado habitualmente como  $\alpha$  o  $\alpha\%$ . Recordemos que el nivel de significación es la probabilidad de rechazar la hipótesis nula ( $H_0$ ), supuesta cierta, es decir,  $\alpha = P(\text{rechazar } H_0 | H_0 \text{ es cierta})$ . Supongamos que trabajamos con un nivel de significación del 5%, esto quiere decir que si  $H_0$  es cierta, se rechazará 5 de cada 100 veces, o lo que es lo mismo, en 100 días que hagamos el control de calidad, y en los cuáles el proceso esté controlado, pararemos el proceso innecesariamente 5 días.

La interpretación errónea más común del nivel de significación consiste en cambiar los dos términos de la probabilidad condicionada en la expresión anterior, es decir, interpretar  $\alpha$  como la probabilidad de que la hipótesis nula sea cierta, una vez que la decisión de rechazarla se ha tomado, esto es, suponer que  $\alpha = P(H_0 \text{ es cierta} | \text{se ha rechazado } H_0)$ . Con respecto a dicho error encontramos por ejemplo las investigaciones de Birnbaum (1982). El autor pidió a sus estudiantes que definiesen el nivel de significación. La respuesta más frecuente fue: “Un nivel de significación del 5% significa que, en promedio, 5 de cada 100 veces que rechazamos la hipótesis nula estaremos equivocados”. En la investigación de Falk (1986) la mayoría de sus estudiantes creían que  $\alpha$  era la probabilidad de equivocarse al rechazar la hipótesis nula. Vallecillos (1994, 1999) también encontró este error en una investigación con más de 400 estudiantes de la Universidad de Granada; y el error se repetía en diferentes tipos de estudiantes (por ejemplo, de medicina, ingeniería o psicología). Resultados similares fueron encontrados por Krauss y Wassner (2002) en profesores de universidad responsables de la enseñanza de métodos de investigación.

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 (1994) describió dos creencias erróneas sobre el tipo de prueba que proporciona el contraste de hipótesis:

1. El contraste de hipótesis es una prueba probabilística de la hipótesis. Permite calcular la probabilidad de que una hipótesis sea cierta. Esta creencia sería cierta sólo cuando aplicamos inferencia bayesiana y además se trataría de una probabilidad subjetiva (personal) o grado de creencia. Cuando trabajamos el método de Fisher o de Neyman y Pearson esta creencia es falsa pues, como hemos dicho, la probabilidad de una hipótesis no tiene sentido en inferencia frecuencial.
2. El contraste estadístico es un método matemático; como tal, y al ser la matemática una ciencia exacta, al finalizar el contraste 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 fe ciega en las matemáticas. 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, 2007), 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.

### **2.5.3. Comprensión del intervalo de confianza**

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 y Fidler, 2005; Cumming, Williams, y Fidler, 2004).

En la investigación llevada a cabo por Cumming y Fidler (2005), los autores mostraron que los intervalos de confianza eran considerados como estadísticos descriptivos, ignorando su naturaleza inferencial. Además comprobaron las dificultades que existen sobre las relaciones que hay entre los distintos conceptos que intervienen en los intervalos de confianza. En otros estudios se pudo comprobar también una cierta debilidad conceptual respecto a la relación existente entre amplitud del intervalo, el tamaño de la muestra y el grado de confianza (Behar, 2001; Fidler, 2005).

## **2.6. CONCLUSIONES**

Las Pruebas de Acceso a la Universidad son un elemento fundamental para aquellos estudiantes que desean continuar sus estudios en la Universidad. Con el fin de poder seleccionar los estudios acordes a sus preferencias será necesario que los contenidos abordados en el curso de segundo de Bachillerato sean los evaluados en dichas pruebas. Por ello es esencial comprobar si exista una correcta correspondencia entre el significado institucional y el evaluado.

En todo proceso de enseñanza-aprendizaje, el conocimiento de los errores más frecuentes cometidos por los estudiantes es un elemento que adquiere gran importancia. Esta información es de gran utilidad para el docente ya que además de saber cómo asimilan y abordan los alumnos un problema, le permitirá llevar a cabo una mejora en la enseñanza. En el caso de la inferencia estadística, esta mejora traerá consigo el perfeccionamiento en sus aplicaciones futuras.

La resolución de un problema de inferencia plantea un reto importante a los estudiantes y determinar claramente cuál es el contenido de los problemas sobre el tema en las Pruebas de Acceso a la Universidad también lo es. Este tipo de análisis adquiere

cierto protagonismo ya que el conocimiento de los contenidos puede promover la obtención de una mejor calificación en dicha prueba y, como consecuencia, facilitar el acceso del estudiante a los estudios universitarios. Además, el estudio y análisis de las PAU será de gran ayuda para el profesor de Bachillerato y al personal encargado de la elaboración de las pruebas.

La información recogida en este capítulo será de gran utilidad con el fin de poder analizar los contenidos de los problemas de inferencia estadística incluidos en las PAU y conseguir así, establecer una idea inicial sobre la dificultad que estas tienen para los estudiantes. De la revisión bibliográfica realizada sobre el análisis de las pruebas de evaluación deja al descubierto la importancia que tiene la presentación de este trabajo con el fin de analizar si las pruebas estas correctamente planteadas para el fin con el que han sido diseñadas.

# Capítulo 3

## ANÁLISIS DE EJERCICIOS DE INFERENCIA ESTADÍSTICA EN LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD

### Índice de Contenidos

---

- 3.1. Introducción
  - 3.2. Metodología
  - 3.3. Muestra de problemas analizados
  - 3.4. Variables analizadas
  - 3.5. Campos de problemas y contenido evaluado en cada uno de ellos
  - 3.6. Modelo probabilístico de la población
  - 3.7. Parámetro a estimar
  - 3.8. Contexto del problema
  - 3.9. Análisis estadístico de los problemas
  - 3.10. Conclusiones sobre la idoneidad didáctica de las PAU
- 

### 3.1. INTRODUCCIÓN

A lo largo del presente Capítulo, se muestra el análisis de los problemas propuestos sobre inferencia estadística llevado a cabo a partir de los ejercicios que han sido propuestos en las Pruebas de Acceso a la Universidad del Distrito Andaluz en la especialidad de Bachillerato de Ciencias Sociales. Concretamente, se han examinado las pruebas de Matemáticas Aplicadas a las Ciencias Sociales II, celebradas desde 2003 hasta 2014.

Para profundizar si los contenidos de inferencia de las pruebas están directamente relacionados con los recogidos en el currículum de Bachillerato de Ciencias Sociales, y el peso relativo que reciben dichos contenidos en el total de la prueba, se analizan los problemas de inferencia formulados en las PAU cuando ha estado vigente el anterior currículum (MEC, 2007). Tal y como se mencionó en el Capítulo 2, los contenidos en la nueva ordenación prácticamente no sufren cambios con respecto a la LOE, aunque se suprime los ejercicios de contraste de hipótesis.

Así pues, el objetivo principal de este capítulo ha sido realizar una caracterización de los problemas relacionados con inferencia estadística incluidos en las citadas pruebas, con el fin de determinar el significado de la inferencia evaluado y, poder así, comparar con el significado pretendido en las orientaciones curriculares. Para ello, mediante un análisis de contenido, se ha definido una serie de variables que han sido consideradas relevantes para la formación del estudiante en relación a los problemas de inferencia y que se han deducido del estudio de la investigación previa. Se han identificado los campos de problemas propuestos, así como el contenido evaluado en cada uno de los campos de problemas. Además, dada las dificultades que encuentran los estudiantes cuando se enfrentan a los ejercicios de inferencia estadística, nos hemos interesado en saber el tipo de variable que está siendo estudiada y, como consecuencia, la distribución de dicha variable; por tal motivo, se han analizado los modelos teóricos de la población de partida y parámetro estimado. Por último, con objeto de valorar la idoneidad emocional y así analizar la conexión entre los conceptos aprendidos y las situaciones de la vida real, se ha examinado el contexto considerando los recogidos en las pruebas PISA (OCDE, 2015).

Dado que el nuevo currículum propuesto en el marco de la LOMCE (MECD, 2015b) 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 y preparar a los alumnos para las pruebas finales de Bachillerato previstas en la LOMCE.

La estructura del capítulo es la siguiente: en primer lugar, se presenta la metodología que ha sido utilizada como base del estudio; a continuación, se describe detalladamente las características de la muestra empleada y cada una de las variables que han sido

objeto de estudio en este trabajo; por último, se finaliza presentando los resultados y conclusiones más destacadas obtenidas del estudio.

### **3.2. METODOLOGÍA**

Krippendorff (2013) señala que el análisis de contenido es una técnica de investigación destinada a formular, a partir ciertos datos, inferencias reproducibles y válidas que puedan aplicarse a su contexto. En la misma línea Weber (1990, p. 9) indica que “el análisis del contenido es un método de la investigación que utiliza un sistema de procedimientos para hacer inferencias válidas del texto”.

En base a estas descripciones, la investigación que se ha llevado a cabo en este trabajo es de tipo cualitativa basada en el análisis de contenido, cuyo foco de atención se centra en el análisis del significado del concepto a través de la descripción de los significados, el abanico de distintos significados y la correcta aplicación del contenido. Con la utilización del análisis de contenido, un texto puede dividirse en unidades que pueden ser clasificadas en un número reducido de categorías en función de variables subyacentes, y que permiten realizar inferencias sobre su contenido (Krippendorff, 2013).

Concretamente, de acuerdo con Bisquerra (1989), el proceso de investigación que se lleva a cabo se enmarca dentro de la tipología inductiva, típico de la metodología cualitativa, donde se pretende analizar los objetos matemáticos implícitos en la descripción de contenidos, criterios de evaluación y estándares de aprendizaje descritos en las Pruebas de Acceso a la Universidad. Así pues, para llevar a cabo la técnica seleccionada, se ha seguido el siguiente plan de actuación (Bisquerra, 1989, p. 62):

1. En primer lugar se ha realizado la recolección de las distintas Pruebas de Acceso a la Universidad en la Comunidad Autónoma de Andalucía en el periodo 2003-2014. La documentación se ha extraído de la página de la Junta de Andalucía, [http://www.juntadeandalucia.es/innovacioncienciayempresa/sguit/g\\_b\\_examenes\\_antteriores.php](http://www.juntadeandalucia.es/innovacioncienciayempresa/sguit/g_b_examenes_antteriores.php).
2. Posteriormente, mediante la resolución de cada uno de los ítems asociados al contenido de inferencia estadística, se han descrito las categorías que

corresponden al ítem para cada una de las variables que son objeto de estudio.

3. Por último, y una vez definidas las categorías y codificados los datos para cada una de las variables, se establecieron regularidades y relaciones entre los datos observados. Un estudio descriptivo de la frecuencia de las variables nos permite obtener conclusiones sobre la presencia global de los mismos, su distribución a lo largo del tiempo y la dificultad de los problemas.

### **3.3. MUESTRA DE PROBLEMAS ANALIZADOS**

Hemos efectuado un estudio longitudinal de análisis de las pruebas propuestas en las pruebas PAU de la modalidad del Bachillerato de Ciencias Sociales en la Comunidad Autónoma andaluza, desde el curso académico 2002-2003 hasta el curso 2013-2014, en la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II. Los contenidos de esta materia son cursados en segundo de Bachillerato en la Modalidad de Humanidades y Ciencias Sociales, y es la continuación de Matemáticas Aplicadas a las Ciencias Sociales I estudiada primero. Esta modalidad de Bachillerato posibilita acceso a estudios universitarios incluidos en la Rama de Arte y Humanidades y la Rama de Ciencias Sociales y Jurídicas.

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. 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 Estadísticas 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 al cuarto ejercicio que corresponde a los ejercicios de contenidos relacionados con la inferencia estadística. En cada uno de los años seleccionados, se revisaron los seis modelos disponibles de prueba (entre los seis

modelos, por sorteo, se elige la prueba de junio y la de septiembre), con dos opciones, cada una de ellas. Cada uno de los modelos que se proponen contiene un ejercicio relacionado con los contenidos de inferencia estadística. Por lo tanto, como cada año contiene 12 relacionados con inferencia estadística, la muestra analizada está compuesta por 144 problemas.

Puesto que la metodología que se está llevando a cabo en este trabajo pretende dar respuesta a problemas concretos, la selección de la muestra se ha realizado de forma intencionada (Goetz y LeCompte, 1988). 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 nuestro estudio pretendemos aportar algunas ideas que pensamos pueden servir para conocer el contenido de inferencia estadística de las PAU, y junto con el análisis y detección de errores, se puede conseguir contribuir a la mejora en el proceso de enseñanza-aprendizaje del tema en estudio. 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.

Como se ha indicado en el Capítulo 2, el análisis realizado en este trabajo se fundamenta sobre las ideas del enfoque Onto-Semiótico (Godino, Batanero y Font, 2007), que asume que los objetos matemáticos surgen de las prácticas (acciones u operaciones) como respuesta a situaciones problemáticas extra o intra matemáticas. En este enfoque, el término “objeto matemático” tiene un significado muy amplio y puede verse como el conjunto de prácticas asociadas a dicho objeto. Mediante un análisis semiótico (Godino y otros, 2007) se ha analizado el contenido matemático, determinando los objetos matemáticos implícitos en dicha resolución.

### **3.4. VARIABLES ANALIZADAS**

Nuestro estudio comenzó clasificando los contenidos recogidos en el Decreto de educación Secundaria y Bachillerato (MEC, 2007), que fueron presentados en Capítulo 1 del presente trabajo. Posteriormente, cada uno de los problemas seleccionados ha sido resuelto con el fin de analizar su contenido matemático.

A partir de estos análisis, se han determinado las siguientes variables, que se describen con más detalle en la sección de resultados.

- **V1. Campo de problema considerado.** Se han diferenciado tres campos principales de problemas: A) muestreo; B) intervalos de confianza; C) contraste de hipótesis. Los dos primeros se han subdividido a su vez en dos subcampos de problema, que serían los siguientes: A1) enumerar las diferentes muestras de una población; A2) identificar la distribución muestral de un estadístico o realizar cálculo de probabilidades con la misma; B1) calcular o interpretar un intervalo de confianza; B2) relación entre confianza, error y tamaño muestral.
- **V2. Modelo probabilístico de la población.** Puesto que la Inferencia Estadística trata de sacar conclusiones de una población en estudio, a partir de la información que proporciona una muestra representativa de la misma, es importante saber las características probabilísticas de dicha población. El conocimiento del modelo probabilístico de la población en estudio nos permitirá saber cómo afrontar la estimación de los parámetros, construcción del intervalo de confianza, formulación de los test de hipótesis, etc. Los modelos o distribuciones de probabilidades analizadas han sido: distribución uniforme discreta, distribución binomial y distribución normal (o distribución gaussiana).
- **V3. Parámetro poblacional que debe ser estimado.** Partiendo del supuesto de que la distribución de la población en estudio es conocida, la inferencia paramétrica se centra en realizar inferencias sobre los parámetros desconocidos del modelo probabilístico. En base a los contenidos recogidos en el Real Decreto de educación Secundaria y Bachillerato (MEC; 2007), los parámetros poblacionales a estimar son: la media, varianza y proporción.
- **V4. Contexto del problema.** Tenemos en cuenta los contextos recogidos en el Programme for International Student Assessment, PISA, organizado por la OCDE (Organización para la Cooperación y el Desarrollo Económico). Las pruebas de evaluación PISA están orientadas a evaluar la competencia matemática, concretamente, evalúan “la capacidad de formular, emplear e interpretar cuestiones matemáticas en diferente tipo de contextos” mediante la resolución de tareas relacionadas con la vida real (MCD, 2013, p. 11). En dichas pruebas participan estudiantes de 15 años de 65 países, incluida España; por lo

tanto, sería deseable que los contextos sugeridos en PISA también se tengan en cuenta en las PAU. Los contextos considerados han sido: personal, profesional, social y científico. Aquellos problemas que no se han enmarcado en alguno de los contextos mencionados, han sido clasificados como problemas sin contexto.

### 3.5. CAMPOS DE PROBLEMAS Y CONTENIDO EVALUADO EN CADA UNO DE ELLOS.

El campo de problemas es un elemento fundamental en el Enfoque Onto-Semiótico (Godino, Batanero y 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. Puesto que el significado del objeto matemático se define a partir del conjunto de prácticas asociadas a los campos de problemas típicos, dichos campos de problemas determinarán en gran medida el significado evaluado en las pruebas. Indirectamente esto influirá en el significado institucional pretendido en la enseñanza de la inferencia en este nivel educativo.

La identificación de los campos de problemas también se considera como parte del análisis de los campos de actuación tal y como pone de manifiesto Rico (2006a). En base a lo comentado, se ha realizado la clasificación de cada uno de los distintos tipos de problemas de inferencia que recogen las pruebas que componen la muestra en estudio. A modo de resumen, la Figura 3.1 recoge los contenidos analizados.

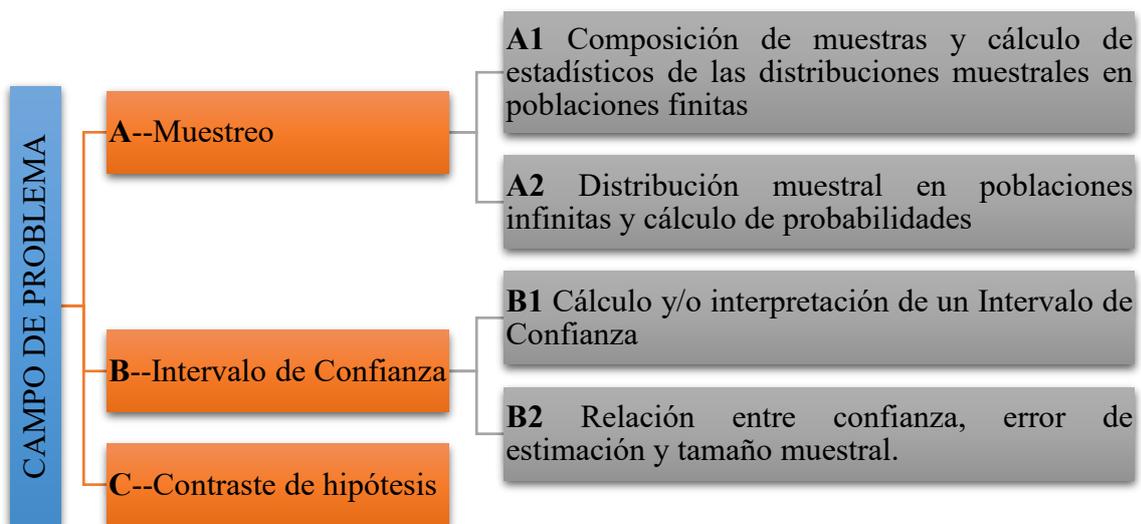


Figura 3.1. Campo de problema evaluado

Con el fin de realizar el análisis semiótico del contenido evaluado y dada la importancia que tiene esta variable, a lo largo de la sección, además de presentar un ejemplo resuelto para cada uno de los campos, se realiza una síntesis del análisis de los campos y se analiza el grado de relación que existe entre el significado pretendido y el significado evaluado para la inferencia en el citado nivel educativo.

### **3.5.1. Análisis semiótico del contenido evaluado**

A lo largo de este apartado se describe con detalle la clasificación sobre los distintos campos de problemas (véase Figura 3.1). De los 144 ítems que forman la muestra, se ha seleccionado un ejemplo correspondiente a cada una de las categorías. Con el fin de analizar la diversidad y especificidad de objetos matemáticos que contiene la resolución del ítem, se han resuelto minuciosamente. Con ello se muestra que la dificultad de los diferentes campos de problemas no es equivalente.

En una primera toma de contacto se ha comprobado que hay ítems que abarcan más de un campo de problema, siendo el número total de ejercicios analizados de 152. Sin embargo, si se realiza el conteo según el número de apartados, este número asciende considerablemente hasta 285.

### **A1. Composición de muestras y cálculo de estadísticos de las distribuciones muestrales en poblaciones finitas**

El objetivo principal de la estadística inferencial es la obtención de conclusiones acerca de una población no observada a partir de una muestra, representativa del total, extraída de dicha población; por tanto representa una colección de métodos para aprender de la inferencia (Harradine, Batanero y Rossman, 2011).

Con objeto de llevar a cabo las generalizaciones deseadas a partir de la muestra seleccionada, se realizan cálculos de algunos resúmenes estadísticos (media muestral, varianza muestral, desviación típica muestral, proporción muestral, etc.), los cuales nos permiten estimar el valor o valores de los parámetros poblacionales (en nuestro caso: media, varianza y proporción poblacional). Recordemos que es imprescindible que las muestras sean aleatorias y sus elementos independientes entre sí con el fin llevar a cabo estas inferencias. Así pues, la selección de los elementos que formen la muestra es un

tema que adquiere gran protagonismo dentro de la inferencia (Chance, del Mas y Garfield, 2004).

Las pruebas PAU contienen ejercicios enfocados en describir los elementos que forman el conjunto seleccionado como muestra y la obtención de las medidas estadísticas en dicha muestra. Por todo lo comentado, dentro de este bloque, nos centramos en estudiar aquellos ítems que incluyen los contenidos mencionados anteriormente para el caso de trabajar con poblaciones finitas. Concretamente, en este apartado se ha analizado los ítems que tratan los siguientes temas:

- composición de las muestras de una población finita en unas condiciones dadas;
- enumeración de las muestras de una población finita en unas condiciones dadas;
- cálculos de sus estadísticos.

Generalmente, los datos que recoge el enunciado del ítem, son: el tipo de población de partida; el tamaño de la muestra y las condiciones del muestreo aleatorio (que puede ser muestreo aleatorio simple, muestreo sistemático y muestreo estratificado). Como ejemplo de este campo de problema, se ha seleccionado el ítem 3B (prueba de reserva) propuesto en el curso académico 2013-2014, cuyo contenido se analiza a continuación.

### **Ítem 3B (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 el primer apartado de este problema, partiendo de una población finita discreta (con tres elementos equiprobables), se extraen aleatoriamente todas las muestras de un tamaño determinado (dos elementos), y se pide calcular para cada una de ellas el estadístico de interés (la varianza muestral). En el segundo apartado se pide la composición de una muestra de tamaño 40 mediante un muestreo aleatorio estratificado con afijación proporcional, facilitando como información el tamaño total de cada estrato.

En relación a la resolución del primer apartado es necesario comprender el tipo de muestreo que se está realizando, pues los resultados obtenidos mediante el muestreo aleatorio con o sin reemplazamiento no coinciden entre sí (las distintas muestras y el número total no coinciden por ambos métodos) (Moore, 2010). 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 en la que los dos elementos que la forman son iguales. Una vez aceptado este hecho, se comienza con la enumeración de todas las posibles muestras de dos elementos, obteniendo como resultado en total nueve muestras diferentes: (6, 6), (6, 9), (6, 12), (9, 6), (9, 9), (9, 12), (12, 6), (12, 9) y (12, 12). Observamos que es necesario tener en cuenta el orden, pues la muestra formada por dos elementos diferentes tiene doble probabilidad que cuando los elementos son iguales. Este es un punto de potencial dificultad, pues los alumnos suelen cometer el llamado error de orden (no tener en cuenta el orden cuando es requerido) (Batanero, Godino, Vallecillos, Green y Holmes, 1994).

Para obtener la varianza de las medias muestrales, se requiere implícitamente el concepto de distribución muestral de la media, el cual se define a través de las medias de cada una de las muestras. Es decir, el alumno debe concebir la media muestral como una variable aleatoria con su distribución, punto que es difícil según Schuyten (1991). Partiendo de dicha distribución muestral, la varianza se obtiene, a partir de la expresión

$$\sigma_{\bar{x}}^2 = \frac{\sum (\bar{x}_i - \mu_{\bar{x}})^2 n_i}{n} \text{ o equivalentemente } \sigma_{\bar{x}}^2 = \frac{\sum \bar{x}_i^2 n_i}{n} - (\mu_{\bar{x}})^2 \text{ siendo } \bar{x}_i \text{ los valores de}$$

las medias de 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 3.1. Haciendo uso de la información recogida en la Tabla 3.1, se tiene que la media y la varianza de la distribución muestral son respectivamente

$$\mu_{\bar{x}} = \frac{\sum \bar{x}_i n_i}{n} = \frac{81}{9} = 9 \text{ y } \sigma_{\bar{x}}^2 = \frac{756}{9} - (9)^2 = 3.$$

De forma alternativa y en base al Teorema Central del Límite es posible obtener la solución del primer apartado. Dicho teorema establece una relación entre la media y desviación de una población no especificada, y la distribución de las medias de las muestras aleatorias con elementos independientes de tamaño  $n$  procedente de dicha población. Concretamente el Teorema Central del Límite nos permite saber que la variable aleatoria media muestral,  $\bar{X}$ , tiene una distribución con media igual a la media

poblacional ( $\mu$ ) y varianza igual a la varianza poblacional dividida por el tamaño muestral ( $\sigma^2/n$ ) (Batanero y Borovcnik, 2016). Por tanto, en nuestro caso, dado que la media y la varianza de poblacional son respectivamente  $\mu = \frac{6+9+12}{3} = 9$  y  $\sigma^2 = \frac{6^2+9^2+12^2}{3} - 9^2 = 6$ , entonces se verifica que  $\mu_{\bar{x}} = 9$  y  $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{6}{2} = 3$ , coincidiendo con los resultados obtenidos anteriormente.

Tabla 3.1.  
*Organización de cálculos en la resolución del problema 3B (2014)*

Medias muestrales ( $\bar{x}_i$ )	$n_i$	$\bar{x}_i n_i$	$\bar{x}_i^2 n_i$
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

Con respecto a la resolución del segundo apartado del ítem 3B, es necesario distinguir los diferentes tipos de muestreo aleatorio que se abordan en los contenidos de la asignatura (muestreo aleatorio simple, estratificado y 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 ( $N=N_1+N_2+\dots+N_k$ , donde  $N_i$  es el tamaño del estrato  $i$ ) y el número total de elementos de las muestra extraída ( $n=n_1+n_2+\dots+n_k$ , siendo  $n_i$  el tamaño de la muestra  $i$  correspondiente al estrato  $i$ ), es decir, se verifica que  $\frac{n_1}{N_1} = \frac{n_2}{N_2} = \dots = \frac{n_k}{N_k} = \frac{n}{N}$ .

En nuestro 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 relación establecida 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 la realización de inferencias respecto a las características de una población (cuyo tamaño excesivamente elevado provoca 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.**

Para cada muestra se verifica que el estadístico puede tomar diferentes valores dependiendo de los elementos que determinan la muestra, por ejemplo, de una misma población se podrían extraer diversas muestras de mismo tamaño pero con medias muestrales distintas. Este hecho, provoca que el estadístico sea considerado como una variable aleatoria con propia distribución, denominada distribución muestral (Moore, 2010). Así pues, la tarea que nos ocupa en este segundo campo de problemas es conocer la distribución de probabilidad de las variables aleatorias asociadas al muestreo o estadísticos muestrales procedentes de poblaciones infinitas. La determinación de la distribución del estadístico muestral nos permitirá realizar inferencia respecto a los parámetros desconocidos de una población. Además, el conocer la distribución muestral de un estadístico nos permitirá hallar probabilidades asociadas al suceso consistente en que el estadístico muestral tome ciertos valores (Behar, 2001).

Dentro de esta categoría de campo de problemas, se han considerado todos aquellos problemas en los que se pide describir la distribución muestral de un cierto estadístico en poblaciones infinitas y el cálculo de probabilidades a partir de dicha distribución muestral, lo que requiere que el estudiante utilice algún modelo teórico de distribución.

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 la distribución binomial a la distribución normal en el caso de muestras de tamaños elevados. Como dato del problema se da, además del tamaño muestral, el tipo de población y el estadístico de interés. En el caso de analizar

la media muestral, en el enunciado se facilita la información referente a la varianza poblacional, mientras que si se trata de la proporción, el estudiante ha de recordar su fórmula (cociente entre el número de veces que se presenta un dato,  $n_i$ , respecto al total de datos que componen la muestra,  $n$ , es decir,  $n_i/n$ ). Se supone que se ha realizado un muestreo aleatorio con reemplazamiento, lo que implica que es posible obtener, como resultado, un número infinito de muestras aleatorias independientes de tamaño  $n$  proveniente de la población de interés. Una vez determinada la distribución del estadístico en estudio, se suele pedir calcular probabilidades, como en el ejemplo que se analiza el ítem 5B propuesto en la convocatoria de septiembre en el curso académico 2010-2011, que resolvemos 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?

En base al Teorema Central del Límite se obtiene, de forma inmediata, la resolución del primer apartado, ya que si una variable aleatoria  $X$  se distribuye según un modelo probabilístico normal con media  $\mu$  y desviación típica  $\sigma$ , entonces, cuando el tamaño muestral tiende a infinito (en nuestro caso,  $n$  superior a 30), entonces se verifica que la variable aleatoria media muestral,  $\bar{X}$ , sigue una distribución  $N\left(\mu, \sigma/\sqrt{n}\right)$ . Por tanto, la resolución del apartado a) consiste en indicar que  $\bar{X} \rightarrow N\left(50, 4/\sqrt{16}\right) = N(50, 1)$ . El alumno ha de recordar estas propiedades.

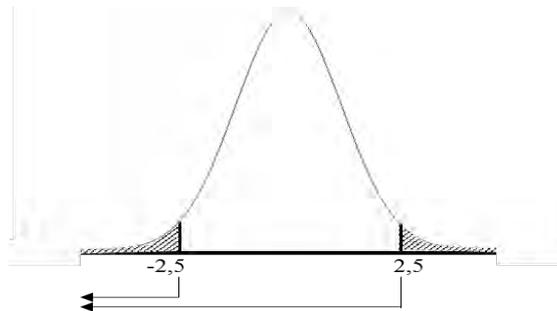
El hecho de describir la distribución de la media de las muestras nos permite obtener probabilidades sobre dicha variable, aplicando los procedimientos de tipificación y lectura de tablas de probabilidades que se conocen. Así pues, tipificando la distribución de la media muestral ( $Z = (\bar{X} - 50)/1$ ), se obtiene la distribución normal tipificada,  $N(0,1)$ , y el estudiante puede consultar las tablas de probabilidades para la resolución del problema (Moore, 2010). A partir de estas tablas y aplicando las reglas simples de cálculo de probabilidades se obtiene:

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

El alumno ha de manejar correctamente las desigualdades para obtener los nuevos límites tipificados. Puesto que la tabla de la distribución normal sólo contiene los valores de probabilidad acumulada hasta un  $Z$  dado, el alumno ha de usar las propiedades aditivas de la probabilidad y de simetría de la curva normal. A partir de la Figura 3.2., se tiene que la  $P\{-2,5 \leq Z \leq 2,5\}$  (zona no rallada en Figura 3.2) se obtiene como:

$$P\{-2,5 \leq Z \leq 2,5\} = P\{Z \leq 2,5\} - P\{Z \leq -2,5\} = 0,9876 .$$

Por lo que, 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 el intervalo delimitado por ambos valores.



*Figura 3.2.* Área delimitada (zona no rallada) por la curva de la distribución normal tipificada,  $N(0,1)$  y las rectas  $x = -2,5$  y  $x = 2,5$

Con el fin de identificar correctamente la distribución de la variable en estudio, el estudiante debe tener en cuenta la relación existente entre la media y varianza de la población y la media y varianza de las muestras de un tamaño determinado. Además, debe ser capaz de tipificar, aplicar reglas básicas de cálculo de probabilidades en un intervalo según una variable aleatoria continua, emplear las reglas de simetría de la curva normal y leer correctamente las tablas de probabilidades de una normal tipificada para conseguir así obtener la probabilidad pedida.

En ambos tipos de problemas el estudiante podrían cometer errores de tipificación o de manejo de las incuaciones requeridas en el cálculo de probabilidades o bien de determinación de percentiles en la curva normal (Batanero, Tauber y Sánchez, 2001).

Los ítems de cálculo de probabilidades tienen un punto de dificultad añadido cuando la variable en estudio es la proporción muestral. De acuerdo a Alvarado y Batanero

(2007) el uso de la corrección de continuidad, requerida para el trabajo de probabilidades de una binomial mediante la aproximación a la normal puede generar cierta confusión. Esta dificultad ha provocado que algunos de los manuales publicados para ser empleados por los estudiantes de segundo de Bachillerato recomienda llevar a cabo este tipo de ejercicios mediante un planteamiento no correcto con el fin de simplificar el proceso (Colera, García y Oliveira, 2003).

### **B1. Cálculo o interpretación de un intervalo de confianza.**

En este bloque se han incluido 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. La construcción de un intervalo de confianza depende de:

- 1) las características de la población, es decir, si el modelo probabilístico es binomial, normal, etc.;
- 2) los parámetros poblacionales desconocidos, como pueden ser por ejemplo la media, varianza y proporción;
- 3) los valores de los estadísticos muestrales conocidos;
- 4) el tamaño muestral.

Sin embargo, aunque a la hora de determinar un intervalo hay que tener en cuenta diversos factores y circunstancias, si bien es cierto, la forma de actuar en su construcción permanece invariable. En lo que sigue analizamos el ítem 1A que fue incluido en la prueba de reserva del curso 2005-2006, siendo la proporción poblacional el parámetro a estimar mediante dicho procedimiento.

#### **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.

La proporción de elementos que poseen una cierta característica en una cierta población es  $p$ . Si se consideran todas las posibles muestras de tamaño  $n$  que pueden extraerse de la mencionada población, se verificará que en cada una de ellas habrá una

proporción de elementos con esa característica. Puesto que este valor varía de una muestra a otra, entonces, la proporción muestral,  $\hat{p}$ , adquiere el carácter de variable aleatoria cuyo modelo probabilístico es una distribución binomial. En base al Teorema Central del Límite y para  $n$  grande ( $n \geq 30$ ) resulta que la distribución binomial se puede aproximar a una distribución normal (Moore, 2010). Por lo que, el estadístico proporción muestral,  $\hat{p}$ , sigue una distribución normal de media  $p$  y desviación típica  $\sqrt{\frac{p(1-p)}{n}}$ , es decir, que  $\hat{p} \rightarrow N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$ . Como consecuencia, al considerar el proceso de tipificación se obtiene que  $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \rightarrow N(0,1)$ .

Un elemento básico en la construcción del intervalo es el coeficiente de confianza o nivel de confianza  $(1-\alpha) \cdot 100\%$ . Dado este coeficiente, se desean encontrar los límites de un intervalo de valores  $(e_1, e_2)$ , tal que se verifique que  $P(e_1 < p < e_2) = 1 - \alpha$ . 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,

$$P\{-k < Z < k\} = 1 - \alpha \tag{3.1}$$

Realizando una lectura inversa de la tabla de la distribución se tiene que, los valores  $-k$  y  $k$  que verifican la expresión (3.1) son aquellos que cumplen que  $P\{Z < -k\} = \alpha / 2$  y  $P\{Z < k\} = 1 - \alpha / 2$ , respectivamente (véase Figura 3.3). A partir de estos valores se despejan los extremos del intervalo de confianza de la variable  $Z$ .

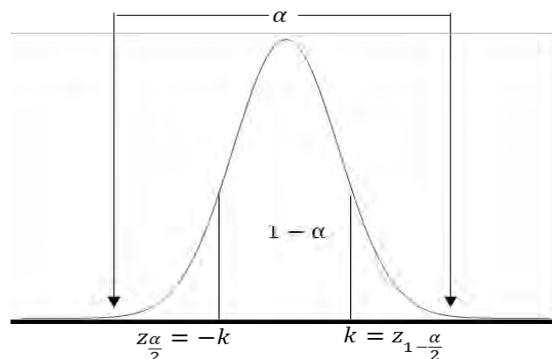


Figura 3.3. Selección de los puntos críticos (para una distribución normal tipificada) en el cálculo del intervalo de confianza

En nuestro caso, puesto que la proporción muestral se distribuye según una normal, tipificando, se obtiene que  $P\left\{|\hat{p} - p| < k\sqrt{\frac{p(1-p)}{n}}\right\} = 1 - \alpha$  y como consecuencia

$$P\left\{\hat{p} - k\sqrt{\frac{p(1-p)}{n}} < p < \hat{p} + k\sqrt{\frac{p(1-p)}{n}}\right\} = 1 - \alpha. \quad (3.2)$$

Una vez despejado el parámetro a estimar, se observa que los límites dependen del parámetro desconocido, véase expresión (3.2). Si  $n$  es suficientemente grande se considera, como una solución satisfactoria, sustituir la proporción poblacional por su estimación. Así pues, en base a los comentarios realizados, se obtiene que el intervalo de confianza de la proporción poblacional, a un nivel de confianza del  $(1 - \alpha) \cdot 100\%$ , viene dada por la expresión<sup>2</sup>

$$I.C._p = \left( \hat{p} - Z_{\frac{1-\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + Z_{\frac{1-\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right).$$

Puesto que  $k = Z_{0,9975} = 2,81$  y  $\hat{p} = \frac{350}{500} = 0,7$  obtenemos que el intervalo de confianza solicitado es  $I.C._p = (0,6421, 0,7576)$ . Un elemento a tener en cuenta y el cual genera gran confusión es la interpretación del resultado obtenido (Cumming y Fidler, 2005; Harradine, Batanero y Rossman, 2011). Una vez determinado el intervalo, no es correcto interpretar que la probabilidad de que el parámetro pertenezca al intervalo es  $(1 - \alpha)$  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 de los  $(1 - \alpha) \cdot 100\%$  que contienen al parámetro, o 0 si el intervalo es uno de los restantes intervalos que no contienen al parámetro. Por esa razón, cuando ya tenemos un valor concreto del intervalo de confianza hablamos de confianza y no de probabilidad, es decir, confiamos en que el intervalo que hemos calculado sea del  $(1 - \alpha) \cdot 100\%$  que contiene el verdadero valor y que en el  $\alpha \cdot 100\%$  restante de ellos el intervalo no lo contendrá (Cumming, Williams y Fidler, 2004).

---

<sup>2</sup> Nótese que  $Z_{\frac{\alpha}{2}} = -Z_{1-\frac{\alpha}{2}}$

## **B2. 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 que conduce al menor error de estimación posible. Puesto que, en general, la expresión del intervalo de confianza depende de la estimación del parámetro, obtenida a partir de la muestra, 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, del parámetro en estudio.

Por todo lo expuesto, hemos visto recomendable analizar dentro de este bloque 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. Un ejemplo de este tipo de ejercicios es el ítem 3B incluido en la prueba de septiembre celebrada en el curso académico 2006-2007.

### **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 es necesario diferenciar entre estadístico (media de la muestra) y parámetro (media de la población) y comprender el significado del intervalo de confianza (rango de posibles valores para el parámetro, en este caso la media de la población) y del nivel 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,  $\bar{X}$ , 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, además de las propiedades anteriores y la definición de intervalo de confianza, que el error de estimación (error máximo admisible o cota de 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 (véase expresión (3.3)). En nuestro caso,

$$E = \frac{1}{2} \times (51,03 - 45,13) = 2,95.$$

$$E = \frac{1}{2} \times \left[ \left( \bar{x} + Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right) - \left( \bar{x} - Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right) \right] \quad (3.3)$$

El segundo apartado, está enfocado en hallar el valor muestral necesario para que la diferencia existente entre la media muestral y la media poblacional sea inferior a 1,8, manteniendo constante tanto la varianza poblacional como el nivel de confianza. Puesto que el objetivo marcado consiste en mejorar la estimación del valor del parámetro disminuyendo el error (manteniendo el nivel de confianza), entonces el único dato a modificar de la expresión  $E = Z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$  será el tamaño muestral.

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 obtenemos la muestra debe tener al menos 267 elementos. El resultado obtenido implicaría que si repitiéramos el experimento, en el 95% de las veces, con un tamaño muestral de 267, se obtendría un error de estimación por debajo a 1,8, mientras que en el 5% restante obtendríamos un error superior.

A partir de la expresión del error de estimación, expresión (3.3), se pueden establecer las siguientes propiedades (Behar, 2001):

- a) a medida que aumenta el tamaño de la muestra (manteniendo constante el nivel de confianza) el error de estimación disminuirá;
- b) manteniendo constante el tamaño de la muestra, el error aumentará conforme aumente el nivel de significación.

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

### C. Contraste de hipótesis

El último tipo de problemas analizado ha sido el contraste de hipótesis. En este punto se solicita al estudiante que compruebe 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 la prueba de reserva de 2013:

#### 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  $H_0: \mu \leq 25$ , 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 una única hipótesis 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, deben determinarse los valores críticos que delimitaran las regiones de aceptación y rechazo en el contraste. En la resolución del problema, será necesario recordar los siguientes conceptos y propiedades (Batanero, 2002; Vallecillos, 1994):

- en la elaboración de un contraste de hipótesis sobre la media de la población, se utiliza como estimador de  $\mu$  la media muestral, cuyo valor debe ser identificando a partir de los datos que facilita el problema;
- si la variable aleatoria  $X$  sigue una distribución normal  $N(\mu, \sigma)$ , entonces para muestras de tamaño 225 se verifica que la variable aleatoria media muestral se distribuye según  $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(\mu, \frac{5}{\sqrt{225}}\right) = N\left(\mu, \frac{1}{3}\right)$ ;
- bajo la suposición de que la hipótesis nula es cierta, es decir, que  $\mu = 25$ , el estadístico de contraste estandarizado,  $Z = \frac{\bar{X} - 25}{1/3}$ , es igual a 3 cuando la media muestral toma el valor 26.

En dicha distribución normal tipificada, el nivel crítico o nivel mínimo de significación, conocido habitualmente como  $p$ -valor, se obtendría como la probabilidad de obtener un valor superior o igual a 3, suponiendo la hipótesis nula  $H_0$  cierta. En nuestro caso se tiene que  $p = P(Z > 3 | H_0 \text{ cierta}) = 0,0013$ . Así pues, si la hipótesis nula fuera cierta, es decir, si el índice medio de masa corporal en su distrito fuese menor o igual a 25, la probabilidad de obtener una muestra de 225 adolescentes con un índice medio igual o mayor a 26 sería 0,0013. Luego, comparando la probabilidad de observar el valor dado ( $p$ -valor) con el nivel de significación prefijado de antemano, se rechazaría la hipótesis nula ya que  $0,0013 < 0,05$  (véase Figura 3.4). Por lo tanto, lo razonable sería rechazar la afirmación del director sanitario ya que los datos empíricos no apoyan la hipótesis nula planteada.

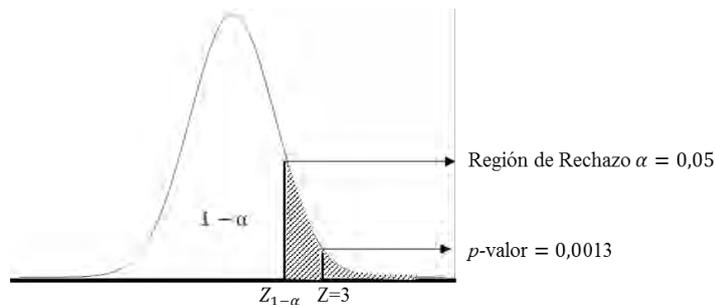


Figura 3.4. Representación de la región de rechazo y  $p$ -valor del contraste de hipótesis

### 3.5.2. Síntesis del análisis de los campos de problemas

De nuestro análisis sobre los problemas de inferencia propuestos en las pruebas PAU, se ha observado el dominio de una gran variedad de conceptos y de procedimientos, como muestra la Tabla 3.2.

Observamos 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): 1) la distribución de los datos en la muestra; 2) la distribución de la variable de interés en la población; 3) la distribución muestral del estadístico en todas las posibles muestras obtenidas de la población.

Tabla 3.2.  
*Contenidos matemáticos evaluados en los diferentes campos de problemas*

		Campos de problemas				
		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 los estadísticos de la población y de la distribución muestral	X	X	X	X	X
	Modelo teórico de distribución	X	X	X	X	X
	Distribución normal tipificada		X	X	X	X
	Estimador de un parámetro			X	X	X
	Intervalo de confianza			X	X	
	Coefficiente de confianza o nivel de confianza; significado			X	X	
	Relación entre amplitud de intervalo, precisión y tamaño muestral				X	
	Hipótesis nula e hipótesis 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
Procedimientos	Enumeración de las muestras de una población finita	X				
	Cálculo de los estadísticos en las muestras	X	X	X	X	X
	Cálculo de los estadísticos de la distribución muestral	X	X	X	X	X
	Cálculo del tamaño de la muestra usando proporciones	X				
	Tipificación		X	X		X
	Lectura de tablas de la distribución normal		X	X	X	X
	Cálculo de probabilidades		X	X		X
	Cálculo de percentiles en la distribución normal			X	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

Esta diferencia (y de los correspondientes resúmenes como la media o la varianza) es compleja y puede llevar al estudiante a cometer errores a la hora de diferenciar, por ejemplo, la varianza de la población y la de la distribución muestral de las medias al resolver los problemas.

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, donde las regiones de aceptación y rechazo se calculan bajo la condición de ser cierta la hipótesis nula (Castro-Sotos, Vanhoof, Noortgate y Onghena, 2007). Puesto que el concepto de probabilidad condicional es difícil y se han descrito números sesgos sobre el mismo (Díaz, Contreras, Batanero y Roa, 2012; Díaz y de la Fuente, 2005; Falk, 1986) los profesores deberán prestar especial atención a su uso en inferencia, para asegurar un aprendizaje adecuado por parte de los estudiantes y garantizar una correcta solución de los problemas.

De la información recogida en la Tabla 3.2, se observa que, puesto que la construcción de intervalos de confianza y el contraste de hipótesis se apoyan sobre los conceptos tratados en A1 y A2 (comprensión del muestreo y distribución muestral), en general, los tres campos, B1, B2 y C, incluyen más contenidos evaluados que los A1 o A2. Esto implica una mayor complejidad de este tipo de problemas respecto a los referidos exclusivamente a muestreo y distribución muestral.

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 al campo C sobre contrastes de hipótesis.

En cuanto a los procedimientos, destacamos la importancia que tiene una correcta lectura la tabla de la distribución normal típica ya que es necesario su uso en todos los campos a excepción de los problemas A1. Señalar también la necesidad de realizar correctamente los cálculos estadísticos con el fin de obtener las medidas de centralización, medidas de dispersión y proporción muestral para posteriormente describir la distribución tanto de la media muestral como de la proporción muestral. Por último señalar que otro punto importante a tener en cuenta es el manejo de desigualdades en el proceso de tipificación.

### **3.5.3. Estudio comparativo entre el significado pretendido y evaluado**

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 la Tabla 3.3, realizamos una descomposición de los contenidos sobre inferencia propuestos para

Matemáticas Aplicadas a las Ciencias Sociales II en el Decreto de Enseñanza Mínimas (MEC, 2007, p. 45476).

En dicha tabla observamos que todos los contenidos fijados en el significado previsto de la inferencia en las directrices curriculares citadas 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 distribución normal serán empleados siempre que se trabaje con la distribución binomial para obtener distribuciones muestrales, calcular o interpretar intervalos de confianza y plantear contrastes de hipótesis referidos a la proporción en una población binomial.

Tabla 3.3.  
*Descomposición del criterio de evaluación sobre inferencia para la asignatura Matemáticas Aplicadas a las Ciencias Sociales II*

	Campos de problemas				
	A1	A2	B1	B2	C
Implicaciones prácticas de los teoremas: Central del Límite, 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	X	X
Distribuciones de probabilidad de las medias y proporciones muestrales.		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: MEC, 2007, p. 45477

Por otro lado, los problemas relacionados con la elección de las muestra sólo se evalúan en los campos de problemas relacionados con ellas, es decir, sobre tipos de muestreo. El conocimiento sobre intervalo de confianza o contraste de hipótesis, sus propiedades y los procedimientos para llevarlos a cabo se evalúan en los campos de problemas B o C respectivamente y el conocimiento de la distribución muestral en todos los campos propuestos.

Aunque el año 2016 es el último previsto para las pruebas de selectividad, se espera que la nueva reválida de Bachillerato, requerida para obtener el título sea muy similar a las actual PAU, pues, además, serán propuestas directamente desde las comunidades autónomas y estarán formadas por problemas abiertos. Es interesante entonces prever si las actuales PAU se ajustarían al nuevo currículo.

La Tabla 3.4 recoge el análisis con respecto a la previsión de aprendizaje de la inferencia, a través de los estándares de aprendizaje evaluables para la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II en el nuevo decreto curricular (MECD, 2015b, p. 389). Se comparan estos contenidos con los evaluados en cada uno de los campos de problemas analizados.

Tabla 3.4.  
*Estándares sobre inferencia para la asignatura Matemáticas Aplicadas a las Ciencias Sociales II*

	Campos de problema				
	A1	A2	B1	B2	C
1.4. Resuelve una situación relacionada con la toma de decisiones en condiciones de incertidumbre en función de la probabilidad de las distintas opciones.					X
2.1. Valora la representatividad de una muestra a partir de su proceso de selección.					
2.2. Calcula estimadores puntuales para la media, varianza, desviación típica y proporción poblacionales, y lo aplica a problemas reales.	X	X	X	X	X
2.3. Calcula probabilidades asociadas a la distribución de la media muestral y de la proporción muestral, aproximándolas por la distribución normal de parámetros adecuados a cada situación, y lo aplica a problemas de situaciones reales.	X	X	X	X	X
2.4. Construye, en contextos reales, un intervalo de confianza para la media poblacional de una distribución normal con desviación típica conocida.			X		
2.5. Construye, en contextos reales, un intervalo de confianza para la media poblacional y para la proporción en el caso de muestras grandes.			X		
2.6. Relaciona el error y la confianza de un intervalo de confianza con el tamaño muestral y calcula cada uno de estos tres elementos conocidos los otros dos y lo aplica en situaciones reales.				X	
3.1. Utiliza las herramientas necesarias para estimar parámetros desconocidos de una población y presentar las inferencias obtenidas mediante un vocabulario y representaciones adecuadas.			X	X	X

Fuente: MECD, 2015b, pp. 389

Podemos observar que el tema de intervalos de confianza tiene mayor peso en estos estándares y el contraste de hipótesis solo se recoge en uno de ellos. También observamos que tanto el cálculo de estimadores puntuales (media, varianza, desviación típica y proporción poblacionales), como el cálculo de probabilidades asociadas a la distribución de la media muestral y de la proporción muestral se recogen en todos los campos de problemas analizados. Nótese que el resto de estándares de aprendizaje son específicos del intervalo de confianza o el contraste de hipótesis. Por último señalar que, aunque la representatividad de la muestra es un elemento de gran importancia dentro de la inferencia estadística, se observa que ninguno de los campos de problemas recoge el estándar de la valoración de la representatividad de la muestra.

### **3.6. MODELO PROBABILÍSTICO DE LA POBLACIÓN DE PARTIDA**

Recordemos que la estadística inductiva o inferencia estadística se centra en la búsqueda de herramientas las cuales nos permiten obtener conclusiones sobre una población a partir del estudio estadístico realizado a una muestra (Batanero y Díaz, 2006). Tanto a la hora de seleccionar una muestra, mediante las distintas técnicas de muestreo, como en el proceso de estimación de la media y la proporción poblacional, se parte del hecho de que la población de partida se distribuye según un modelo teórico determinado. Por tanto, a la hora de comenzar el estudio es necesario tener bien claro el concepto de población y las características de ésta, es decir, será necesario tenerla perfectamente definida (Rossman, 2008).

Dada la diversidad de modelos probabilísticos que tienen juego en los problemas propuestos en las PAU, se ha analizado la distribución o modelo probabilístico de la población de partida.

En la realización de un muestreo aleatorio *a priori* se supone que la probabilidad de ser elegidos es la misma para todos y consecuentemente la población de donde ha sido extraída dicha muestra se distribuye según una uniforme discreta. Ejemplo de ello son los dos ejercicios incluidos en el ítem propuesto en la prueba de reserva 2B del curso 2013-2014. En el primero no se especifica el tamaño de la población, pero hemos de suponer que todos los alumnos tienen la misma probabilidad de entrar en la muestra; por

tanto la distribución es uniforme. En el apartado b) se trata de una pequeña población con tan sólo cinco elementos.

**Ítem 2B (2014):**

- 1) En un centro docente la tercera parte de los alumnos estudia el idioma A, la mitad el idioma B y el resto el idioma C (cada alumno estudia sólo uno de estos idiomas).
  - a) Se desea seleccionar una muestra de 60 alumnos, mediante muestreo aleatorio estratificado con afijación proporcional al número de los alumnos de cada idioma. ¿Cómo debería estar conformada la muestra?
  - b) En otra muestra seleccionada por el procedimiento anterior, el número de alumnos tomados del idioma A es 14. Determine cuántos se han elegido de los otros dos idiomas.
- 2) Una población tiene 5 elementos. Mediante muestreo aleatorio simple se seleccionan muestras de tamaño 3, siendo la desviación típica de sus medias 2 y la media de las medias muestrales 7. ¿Cuánto valen la media y la varianza de la población?

En el caso de realizar alguna inferencia sobre el valor de la media poblacional se supone que la población sigue una distribución normal con desviación típica conocida, que se da como dato del problema. En esta situación se verifica que si la población de partida es normal con varianza conocida entonces la distribución muestral de las medias es también normal. En el ejemplo ítem 5A de 2009 se facilita al estudiante la información referente a la distribución de la población en estudio que es necesario en la construcción del intervalo de confianza.

**Ítem 5A (2009):**

- Una variable aleatoria  $X$  se distribuye de forma Normal, con media  $\mu$  y desviación típica  $\sigma=0,9$ .
- a. Una muestra aleatoria de tamaño 9 ha proporcionado los siguientes valores de  $X$ : 7,0; 6,4; 8,0; 7,1; 7,3; 7,4; 5,6; 8,8; 7,2. Obtenga un intervalo de confianza para la media  $\mu$ , con un nivel de confianza del 97%.
  - b. Con otra muestra, se ha obtenido que un intervalo de confianza para  $\mu$ , al 95%, es el siguiente (6,906, 7,494). ¿Cuál es el tamaño de la muestra utilizada?

En el caso de los problemas relaciones con la estimación de la proporción poblacional se supone que la distribución de la población de partida es una distribución binomial. Este tipo de problema suele revestir mayor dificultad, puesto que, por un lado, no se da el valor de la varianza o desviación típica y es el alumno quien debe recordar la fórmula para determinarla, y por otro lado, puesto que las muestras suelen ser grandes, hay que utilizar la aproximación de la distribución binomial a la distribución normal. Ejemplo de ello es el ítem 6A (2011) en el que el alumno debe identificar la distribución de la población ya que no es facilitada como dato.

**Ítem 6A (2011)**

El director de una televisión afirma que un nuevo programa que va a emitirse será visto, al menos, por un 30% de personas. Una vez emitido se realizó una encuesta a 500 personas, elegidas al azar, y ésta reveló que 130 de ellas habían visto ese programa.

- a) Formule la hipótesis nula y la alternativa del contraste de hipótesis que permite determinar si los datos de la encuesta realizada son compatibles con la afirmación del director.
- b) Halle la región crítica de ese contraste para un nivel de significación del 5.5%.
- c) Según el dato obtenido en el apartado anterior ¿qué conclusión se obtiene sobre la afirmación realizada por el director de esa televisión?

Realizando el estudio a partir de los campos de problemas se observa que a excepción del campo A1, donde la distribución de la población es uniforme discreta, en el resto de los campos (distribución muestral en poblaciones infinitas, cálculo de probabilidades, construcción e interpretación de intervalos de confianza y contraste de hipótesis) los modelos probabilísticos de la población son binomial o normal, véase Tabla 3.5.

Tabla 3.5.  
*Modelos probabilísticos incluidos en los campos de problema*

	Campos de problema				
	A1	A2	B1	B2	C
Distribución Uniforme Discreta	X				
Distribución Binomial		X	X	X	X
Distribución Normal		X	X	X	X

Fuente: Elaboración propia

Por último señalar que, aunque todos los campos, a excepción del A1, trabajan tanto con la distribución normal como con la binomial, en general, y tal como se ha puesto de manifiesto en el análisis sobre los procedimientos recogidos en la Tabla 3.2, en un alto porcentaje de ejercicios se termina recurriendo al uso de la tabla de la distribución normal. Por lo que se puede se pone de manifiesto la importancia que tiene el uso distribución gaussiana o distribución normal dentro de la inferencia estadística.

### 3.7. PARÁMETRO POBLACIONAL OBJETO DE ESTUDIO

En este apartado se han analizado los parámetros que deben ser estimados mediante estimación puntual (a partir de los estadísticos muestrales, se asigna un valor al parámetro poblacional), estimación por intervalo de confianza (intervalo dentro del cual está contenido el verdadero valor del parámetro desconocido con cierto grado de certeza) y aquellos sobre los que se plantea una hipótesis a contrastar (mediante el

desarrollo de un procedimiento establecido se comprueba la verosimilitud de la hipótesis planteada). No todos los problemas solicitan la estimación del mismo parámetro o bien la realización de un contraste sobre el mismo y, aunque no siempre, generalmente el tipo de parámetro que se pide se relaciona con la distribución en la población.

En general los parámetros a estimar, analizando cada uno de los campos de problemas (véase Tabla 3.6) son los referentes a la media poblacional y la proporción poblacional.

Tabla 3.6.  
*Clasificación de los parámetros a estimar según los campos de problema*

	Campos de problema				
	A1	A2	B1	B2	C
Proporción	X	X	X	X	X
Media	X	X	X	X	X
Varianza/Desviación típica	X				

Con respecto a la estimación de la varianza poblacional, solamente es posible encontrarla en los ítems asociados a la categoría A1 (composición de las muestras y cálculos estadísticos de las distribuciones muestrales en poblaciones finitas) ya que en todos los relacionados con la estimación de media poblacional siempre se parte de que dicho parámetro es conocido. A continuación se muestra un ejemplo para cada uno de los tres tipos de parámetros diferentes: varianza poblacional (ítem 3A curso 2003-2004) media poblacional (ítem 4A curso 2004-2005) y proporción poblacional (ítem 3A curso 2005-2006).

#### **Ítem 3A (2004)**

Dada la población de elementos  $\{3, 4, 5, 8\}$ , se pretende seleccionar una muestra de tamaño 2, mediante muestreo aleatorio con reemplazamiento.

- Escriba todas las muestras posibles.
- Calcule la varianza de la población.
- Calcule la varianza de las medias muestrales.

#### **Ítem 4A (2005)**

La superficie de las parcelas de una determinada provincia se distribuye según una ley Normal con media 2,9 Ha y desviación típica 0,6 Ha.

- Indique la distribución de las medias muestrales para muestras de tamaño 169.
- ¿Cuál es la probabilidad de que una muestra de tamaño 169 tenga una superficie media comprendida entre 2,8 y 3 Ha?

### **Ítem 3B (2006)**

Se ha lanzado un dado 400 veces y se ha obtenido 80 veces el valor cinco. Estime, mediante un intervalo de confianza al 95 %, el valor de la probabilidad de obtener un cinco.

## **3.8. CONTEXTO DEL PROBLEMA**

El contexto en el que se enmarca un problema es esencial para mostrar la unión existente entre los conceptos y las situaciones reales, permitiendo además darle sentido al aprendizaje y motivar el interés por el aprendizaje de las matemáticas, haciéndoles ver su utilidad en diferentes contextos. Para tal fin, nos basaremos en los contextos recogidos en las pruebas de evaluación PISA orientadas a evaluar la competencia matemática.

El análisis de este elemento adquiere cierta importancia ya que tal y como señala Rico (2006b) el uso de las matemáticas en situaciones y contextos variados es un componente importante de la alfabetización matemática. Las situaciones permiten conectar un problema con los fenómenos de los que surgen los objetos matemáticos considerados. A continuación se presenta cada uno de los contextos junto a un ejemplo extraído de las pruebas PAU.

- *Situación personal:* Son problemas relacionados con las actividades del día a día del propio individuo, su familia y su grupo de iguales. Como ejemplo, se muestra el ítem 3B propuesto en 2004, pues el contexto es el número de horas que dedica el estudiante a la práctica de deporte, que es una variable personal para el alumno. Otros ejemplos de problemas en contexto personal son los relacionados con las compras, familia, juegos, salud o transporte personal, deportes, o viajes.

### **Ítem 3B (2004)**

El número de horas semanales que los estudiantes de Bachillerato de una ciudad dedican al deporte se distribuye según una ley Normal de media 8 y varianza 7,29.

- a) Para muestras de tamaño 36, indique cuál es la distribución de las medias muestrales.
- b) ¿Cuál es la probabilidad de que la media de una muestra de tamaño 36 esté comprendida entre 7,82 y 8,36 horas?

- *Situación profesional:* Son problemas que se centran en el mundo laboral que el estudiante encontrará en el futuro o que conoce por sus padres o maestros. Se trata de problemas sobre medida, control o coste de un proceso de producción o

una construcción, sobre diseño en carpintería, arquitectura o jardinería; coste o salario de mano de obra, etc. En nuestro caso se ha considerado como ejemplo el ítem 1A, propuesto en 2007, donde la variable aleatoria que se está estudiando es el salario de los trabajadores de una ciudad.

### **Ítem 1A (2007)**

El salario de los trabajadores de una ciudad sigue una distribución Normal con desviación típica 15 euros. Se quiere calcular un intervalo de confianza para el salario medio con un nivel de confianza del 98%. Determine cuál es el tamaño mínimo de la muestra que se necesitaría recoger para que el intervalo de confianza tenga una amplitud, como máximo, de 6 euros.

- *Situación Social:* En este caso, se incluyen los problemas que el estudiante podría encontrar en su comunidad más amplia que la familiar (comunidad de vecinos, ayuntamiento o ciudad, su país, etc.). El ítem 1B de 2012 se ha incluido dentro de esta categoría por tratar el tema del uso del carril bici en una ciudad. Otros casos encontrados son problemas relacionados con sistemas electorales, transporte urbano o interurbano, demografía o publicidad. Dentro de este tipo se han incluido también los juegos de azar como la lotería, etc., que aparecen en los medios de comunicación.

### **Ítem 1B (2012)**

Un informe de un Ayuntamiento afirma que al menos el 26% de los usuarios del carril bici habrían utilizado el coche particular para sus desplazamientos de no haber existido dicho carril. Sin embargo, un periódico local anuncia la falsedad del dato, informando que una encuesta propia indica que solo 240 de los 1000 usuarios encuestados afirman que habrían utilizado el coche particular.

- a) Establezca un contraste, con hipótesis nula, para verificar la afirmación del Ayuntamiento e indique la región crítica de dicho contraste para un nivel de significación del 5%.
- b) Con este nivel de significación, ¿podría aceptarse el informe del Ayuntamiento?

- *Situación Científica:* Los problemas clasificados en la categoría científica están relacionados con la aplicación de las matemáticas en ciencia y tecnología. A modo de ejemplo se muestra el ítem 6A propuesto en 2013, cuyo objetivo es realizar un estudio sobre la proporción de hembras entre los peces de una piscifactoría. Otros problemas relacionados con la ciencia tratan de la meteorología, ecología, medicina, genética, o física. Esta situación es más abstracta que el resto, ya que implica la comprensión de un proceso tecnológico, una interpretación teórica o un problema matemático.

### **Ítem 6A (2013)**

Se quiere estimar la proporción de hembras entre los peces de una piscifactoría; para ello se ha tomado una muestra aleatoria de 500 peces, y en ella hay 175 hembras.

- a) Calcule un intervalo de confianza para la proporción de hembras en esta población de peces, con un nivel de confianza del 94%.
- b) A la vista del resultado del muestreo se quiere repetir la experiencia para conseguir un intervalo de confianza con el mismo nivel y un error máximo de 0,02, ¿cuál es el tamaño mínimo que debe tener la nueva muestra?

Por otro lado, aquellos problemas que no han sido incluidos en ninguno de los contextos mencionados anteriormente, han sido clasificados como problemas *Sin contexto*. Concretamente, se han incluido en esta categoría aquellos enunciados redactados sobre situaciones matemáticas abstractas y en los que no se incluye ninguna aplicación a la vida real. Ejemplo de ello es el ítem 5A propuesto en 2012. Señalar que la inclusión de este tipo de ejercicios contradice todas las recomendaciones sobre la enseñanza de la estadística en que el contexto es fundamental para el aprendizaje.

#### **Ítem 5A (2012)**

Una característica de una determinada población se distribuye según una variable aleatoria Normal  $X$  de media desconocida y desviación típica 0,9. Extraída al azar una muestra de tamaño 9 de esa población y observada  $X$ , dio como resultados: 10,5; 10; 8,5; 10,5; 11,5; 13,5; 9,5; 13; 12.

- a. Halle un intervalo de confianza, al 99%, para la media de la variable  $X$ .
- b. Determinar el tamaño mínimo que debe tener una muestra de esa población, para que el error máximo que se cometa en la determinación de un intervalo de confianza para la media de  $X$  sea, a lo sumo, 0,3, con un nivel de confianza del 90%.

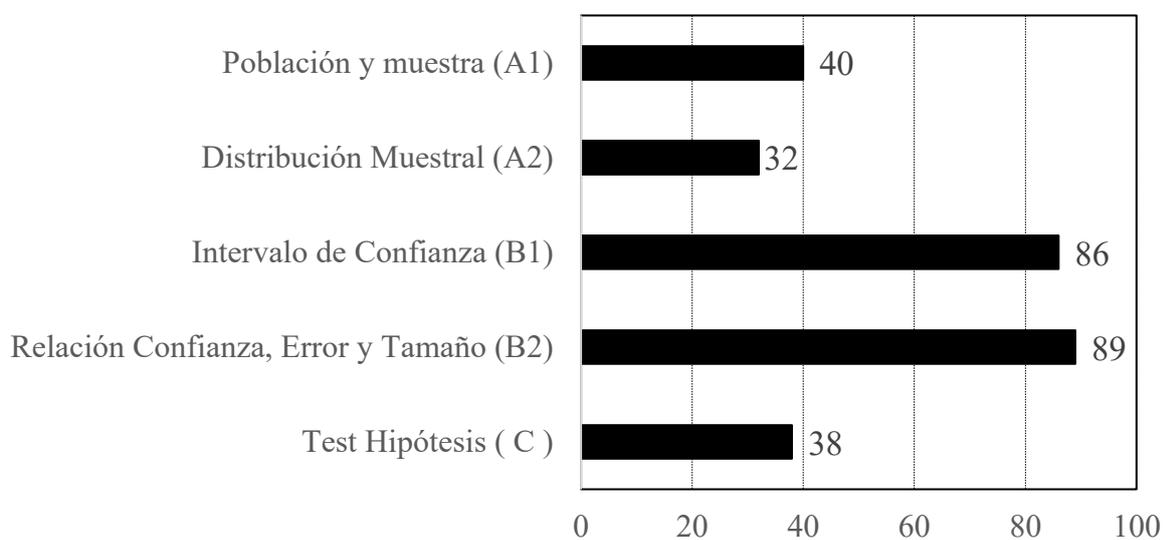
### **3.9. ANÁLISIS ESTADÍSTICO DE LOS PROBLEMAS**

A partir del análisis cualitativo, donde se identificaron las categorías para cada una de las variables y para cada problema, se han codificado los datos para su posterior tratamiento estadístico con el software Excel. En lo que sigue, se presenta la distribución, globalmente y por año, de las categorías en la muestra de problemas analizados. Los resultados serán presentados siguiendo el mismo orden de análisis realizado para cada una de las variables.

#### **3.9.1. Campos de problemas**

Tal y como se ha mencionado en la Sección 3.5 los campos de problemas que se han tratado son: A) muestreo (dividido en las categorías A1 y A2); B) intervalos de confianza (dividido en las categorías B1 y B2); C) contrastes de Hipótesis. En la Figura

3.5 se muestra la distribución global de los campos de problemas propuestos en los 144 ejercicios analizados. Considerando los tres campos de problemas, se ha comprobado que en un mismo ejercicio es posible encontrar ejercicios que tratan más de un campo, por lo que el total de ejercicios analizados se incrementa a 152 (35 correspondientes al campo A; 95 al campo B y 22 al campo C). Si se analizan de forma independiente los apartados que se incluyen en las 144 pruebas, se obtiene que el número total de la muestra casi duplica el número inicial (285). Este hecho nos indica que 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.



*Figura 3.5.* Frecuencia de apartados según campo de problema

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 175, por lo que se deduce la gran importancia dada a este tipo de ejercicios 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, hay un buen ajuste entre los problemas propuestos en las pruebas de acceso y los estándares de evaluación previstos en el nuevo desarrollo

curricular (MECD, 2015b). Sin embargo, no se contempla adecuadamente el contraste de hipótesis, previsto en el desarrollo curricular actual (MEC, 2007).

Por otro lado, si se analizan los distintos tipos de ítems propuestos en cada uno de los años (véase Figura 3.6) se observa que no es hasta el curso académico 2009-2010 donde se pueden encontrar pruebas que incluyan ejercicios relacionados con test o contrastes de hipótesis. A partir de su inclusión en las PAU, casi la mitad de los ejercicios propuestos lo incluyen. La aparición de los contrastes de hipótesis está suponiendo una disminución considerable de los ejercicios relacionados con el intervalo de confianza, indicando un cambio de tendencia en los contenidos de inferencia evaluados en las pruebas de acceso. A diferencia de los otros campos de problemas, se ha detectado que la mitad de los ejercicios relacionados con test de hipótesis no están subdivididos en apartados (11 de los 22 ejercicios), este hecho acentúa la dificultad de este tipo de ejercicios ya que el estudiante se juega toda la calificación (2,5 puntos) a un único apartado. Además, consideramos que este cambio aumenta la dificultad pues autores como Fidler y Cumming (2005) indican que la comprensión de los intervalos es más sencilla que la de los contrastes tanto para estudiantes como para investigadores. Además contradice el nuevo currículo en el que se ha suprimido el contraste de hipótesis.

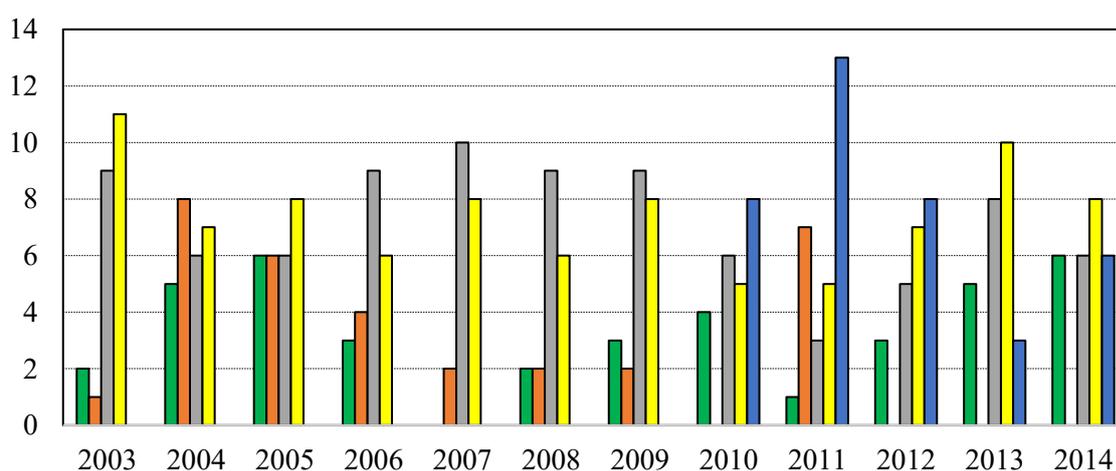


Figura 3.6. Clasificación por años según los campos Población y Muestra, (A1, verde); Distribución Muestral (A2, naranja); Intervalo de confianza (B1, gris); Relación entre confianza, error y tamaño (B2, amarillo); y Contraste de Hipótesis (C, azul)

Finalmente resaltamos el hecho de que los problemas relacionados simplemente con muestreo (categorías A1 o/y A2) tienen una distribución muy variable. Aunque los encontramos en todos los años, cabe destacar que los problemas donde se solicita la descripción de la distribución muestral junto al cálculo de probabilidades han sido incluidos en ocho de los doce años analizados, y además, a partir de 2010, y con excepción del año 2011, no han sido recogidos en las PAU.

### 3.9.2. Modelo probabilístico de la población de partida

El análisis de los modelos probabilísticos que sigue la población en estudio nos muestra la importancia que tiene dentro de la inferencia estadística la distribución normal. Dado que hay ítems que contienen más de un apartado, se tiene que de los 155 apartados analizados se observa que la población de partida se distribuye según una normal en 91 de ellos (58%), véase Figura 3.7. Por otro lado, aunque generalmente a posteriori se suele realizar la aproximación de la distribución binomial a la normal, los problemas relacionados con proporciones nos han permitido concluir que el 26% de ellos parten de una población binomial. En otros casos, la población se distribuye según una uniforme de tipo discreto.

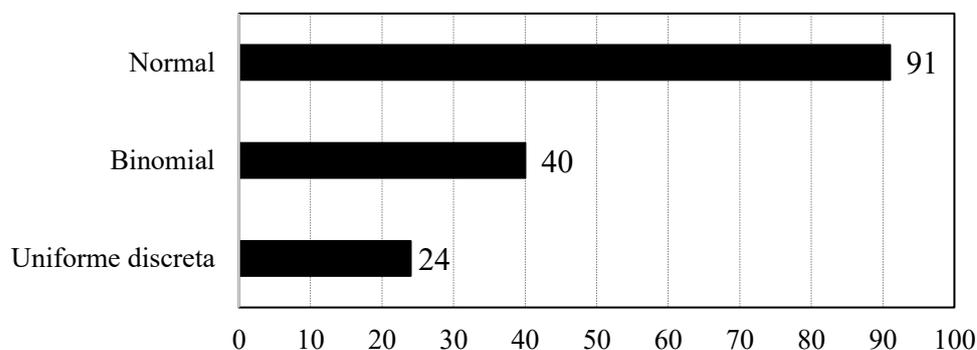


Figura 3.7. Frecuencia de apartados según modelo probabilístico de la población de partida

Realizando la clasificación por años, véase Figura 3.8, y comparándolo con los resultados obtenidos en el punto anterior, se comprueba que el hecho de no incluir ítems de la categoría A1 en 2007 provoca que no tengamos poblaciones uniformes discretas. A partir de la información recogida en la Figura 3.8, se puede concluir que no es hasta el 2006 donde se introducen ejercicios donde el parámetro en estudio sea la proporción

muestral, ya que en años previos no se ha contabilizado ningún ítem donde la población se distribuya según una binomial. La introducción de la distribución binomial en las pruebas ha provocado una disminución en los ítems que consideran la distribución normal como distribución de partida de la población.

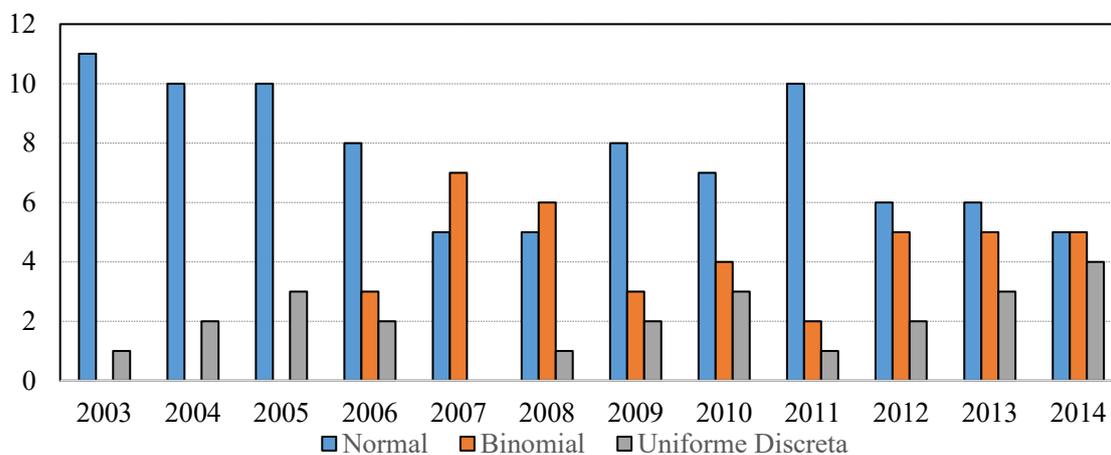


Figura 3.8. Clasificación de los modelos probabilísticos de la población por años

### 3.9.3. Parámetro poblacional objeto de estudio

Dentro del campo de la inferencia estadística paramétrica, en general, se supone que la variable aleatoria en estudio tiene un modelo de distribución de probabilidad donde se desconoce alguno de los parámetros que la determinan y cuyo objetivo está centrado en la estimación del mismo. Por todo lo comentado, es de esperar que los resultados obtenidos en las subsecciones 3.9.1 y 3.9.2 tengan una alta relación con los que se muestran a continuación.

La Figura 3.9 recoge los resultados obtenidos después de analizar el tipo de parámetro que es objeto de estudio en el ítem. Como se observa, el número total de parámetros a estimar (puntualmente, por intervalo de confianza o mediante la realización de un test) es 120, siendo muy inferior al número total de campos de problemas clasificados (285). Recordemos que a la hora de realizar la clasificación no se ha tenido en cuenta los ejercicios en los que se solicita la obtención de la distribución de la media muestral mediante el uso del Teorema Central del Límite y el posterior cálculo de una probabilidad determinada.

A partir de la Figura 3.9 se observa que la media poblacional es el parámetro a estimar que tiene mayor peso en las pruebas PAU (65%). En casi la mitad de los casos

encontramos que el parámetro que se desea estimar es la proporción poblacional. Teniendo en cuenta los resultados mostrados en la subsección 3.9.2., este hecho era de esperar ya que la introducción de este parámetro se produce a partir de 2006. Señalamos el hecho de haber encontrado únicamente dos ítems donde el parámetro a estimar es la varianza. Concretamente estos dos ejercicios son el ítem 3A propuesto en el curso 2003-2004 presentado en la Sección 3.7 y el ítem 2B propuesto en el curso académico siguiente.

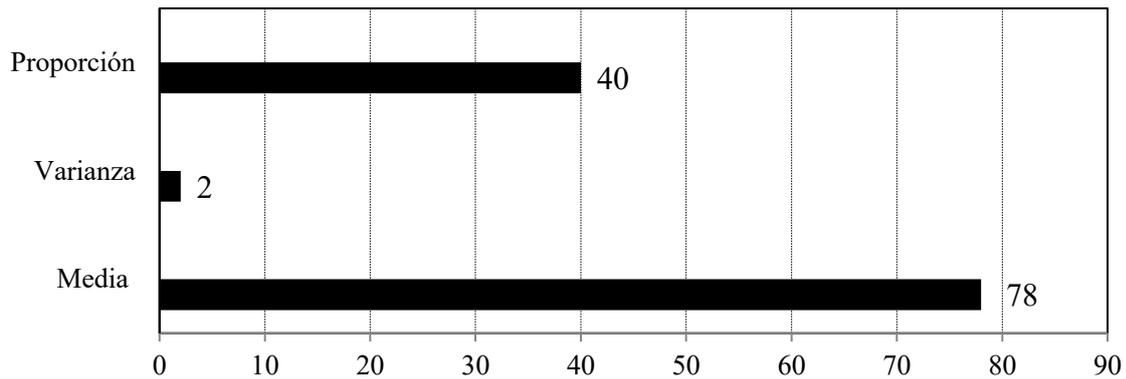


Figura 3.9. Frecuencia de los parámetros a estimar

En el análisis por año, véase Figura 3.10, se observa como la proporción poblacional va adquiriendo mayor protagonismo llegando a tener el mismo peso en el último año analizado.

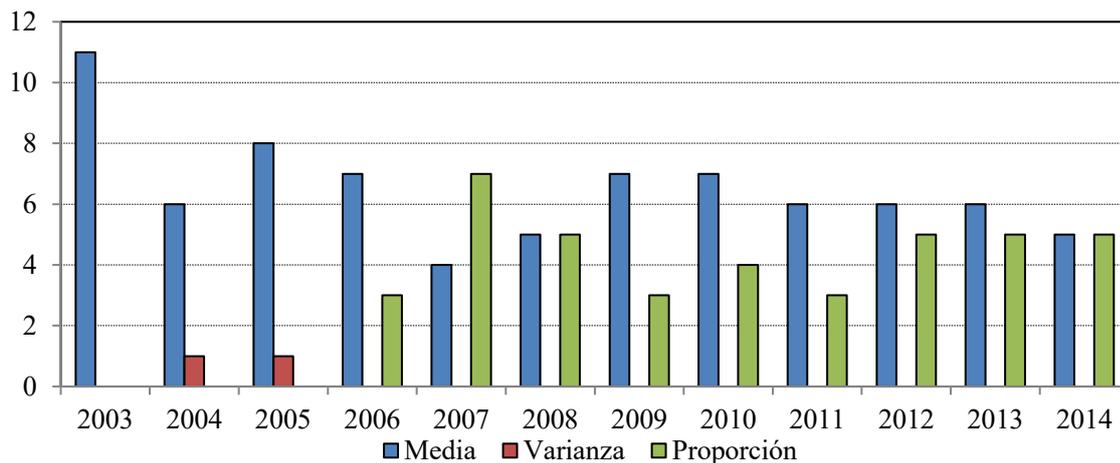


Figura 3.10. Clasificación de los parámetros a estimar por años

### 3.9.4. Contexto

Partiendo de la información recogida en la Sección 3.8, se han analizado los contextos para cada una de las pruebas. Con objeto de tener una visión más general sobre la distribución de contextos utilizados se presenta la Figura 3.11. Queremos destacar que el número total de contextos analizados (153) supera al número total de ítems (144) ya que algunos de los problemas propuestos incluían varios apartados independientes entre sí y enmarcados en contextos distintos. Al realizar la clasificación se observa que se tiende a proponer problemas dentro de un contexto social, ya que aproximadamente el 29% de ellos se clasifican como tales.

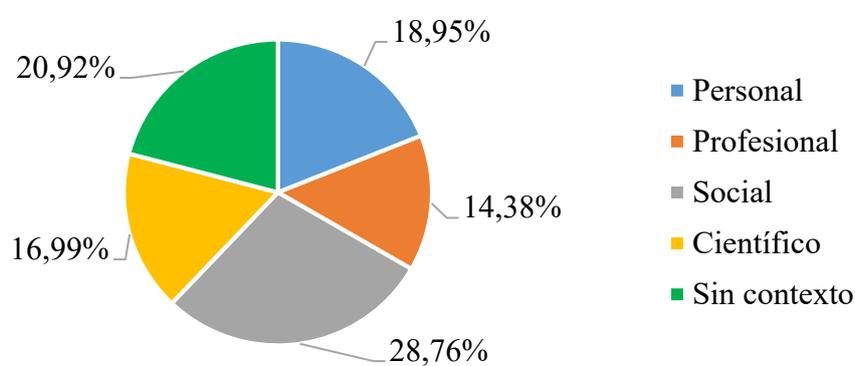


Figura 3.11. Clasificación según contexto.

Aunque a partir de la información recogida en la Figura 3.11 se concluye que aproximadamente el 21% de los problemas analizados carecen de contexto, es necesario señalar que el 75% de ellos han sido propuestos antes de 2010. Este hecho adquiere cierta importancia ya que se tiende a seguir las recomendaciones sobre la enseñanza de la estadística disminuyendo los problemas sin contexto donde el estudiante no comprende el interés de aplicación de las matemáticas.

Si analizamos los contextos por año (véase Figura 3.12), se observa una gran presencia de problemas formulados sobre un contexto social y una disminución significativa, en los últimos años, de los problemas sin contexto. Al realizar un análisis comparativo entre el periodo inicial y final del estudio, se observa que ha habido una tendencia a modificar el contexto en el que se desarrolla el problema. Por ejemplo en 2003 se tiene una alta presencia de problemas con contextos propiamente científicos y

no aparecen problemas con contextos personales, mientras que en 2014 se tiene que más del 50% de los problemas corresponden a un contexto social y no es posible encontrar problemas propios de un contexto científico.

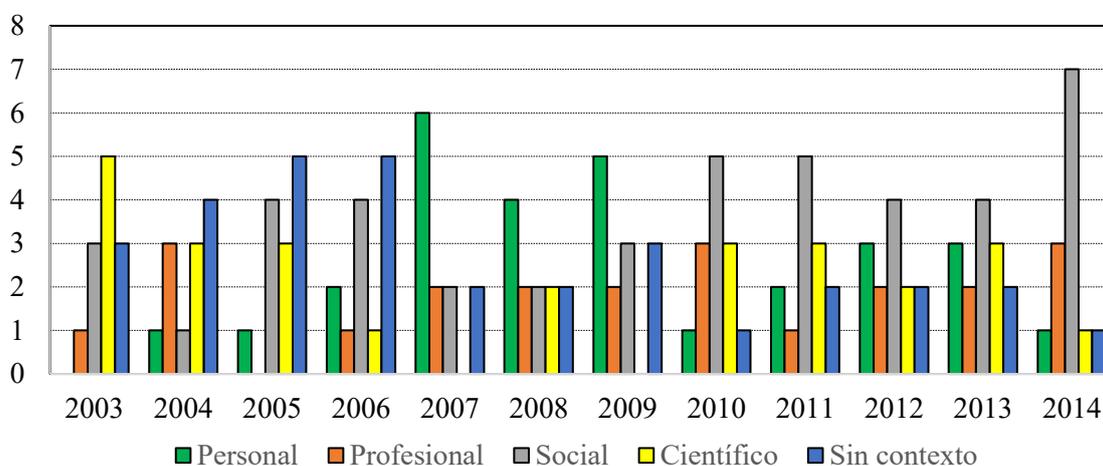


Figura 3.12. Contextos de los ítems por cada año

### 3. 10. CONCLUSIONES SOBRE LA IDONEIDAD DIDACTICA DE LAS PAU

Para finalizar el Capítulo hacemos una reflexión sobre algunos componentes de la idoneidad didáctica de las PAU que se pueden deducir del análisis llevado a cabo.

- *Idoneidad epistémica.*

Mediría la correspondencia entre el significado institucional pretendido (dado por las directrices curriculares) y el evaluado. Los resultados del análisis realizado muestran que esta idoneidad es escasa. Por un lado, todas las pruebas de los catorce años analizados incluyen un problema de inferencia estadística. Aunque el contenido del currículo, incluye otros muchos temas relacionados con inferencia estadística objetivamente se ve que las pruebas dan mucha mayor importancia al cálculo intervalos de confianza y la interpretación de los mismos, que a otro temas de inferencia pues aproximadamente el 64% de los ítems propuestos van relacionados con este tema. Sin embargo, vemos que no todos los temas de inferencia se contemplan con el mismo peso. Por otro lado, mientras el currículo contempla tres bloques de contenidos -Álgebra, Análisis y Probabilidad y Estadística- la mitad de las pruebas están dedicadas a este

último bloque. En consecuencia, no hay una correspondencia adecuada entre el significado pretendido y evaluado.

- *Idoneidad cognitiva.*

Mide el grado en el que los significados evaluados en las pruebas son asequibles a los estudiantes. Pensamos que en este caso esta idoneidad es también baja coincidiendo con algunos de los resultados de investigaciones que analizan las calificaciones de los estudiantes (Grau, Cuxart, y Martí-Recober, 2002; Murillo, 1997; Ruiz, Dávila, Etxeberría y Sarasua, 2013)

Por otro lado, analizando la solución de alguno de estos problemas, hemos visto que los problemas propuestos son bastante complejos, pues incluyen gran cantidad de objetos matemáticos. Otro indicio de que se quiere aumentar la dificultad es la proporción de ejercicios de test de hipótesis que se están planteando en los últimos años los cuales requieren el uso de muchos conceptos: población, muestra, parámetro poblacional, estadístico muestral, distribución del estadístico muestral, hipótesis nula y alternativa, región de aceptación y rechazo, nivel de significación.

Además, señalar que 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. Esta enseñanza soslaya también la problemática filosófica asociada y los errores de aplicación de la inferencia. Además, presenta la inferencia frecuencial como una metodología única, ocultando las diferentes aproximaciones y las controversias que dentro de la misma estadística ha tenido la inferencia (Batanero, 2000).

- *Idoneidad mediacional.*

Esta idoneidad mide la adecuación de los recursos materiales y temporales que se proporcionan al estudiante para la realización de las pruebas. Fundamentalmente será el tiempo disponible, las tablas de la distribución normal y las calculadoras. Pensamos que estos recursos son suficientes, sin embargo, es paradójico el hecho de que tanto la estadística como la probabilidad con frecuencia se dejan en los cursos anteriores como último tema y a veces se omiten; 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.

- *Idoneidad emocional.*

Mide el interés o motivación que puede tener el estudiante con el objeto de estudio. En nuestro caso, al analizar los contextos que se presentan en los problemas, se ha observado un predominio del contexto social que es el que en principio debe tener más interés para los estudiantes de esta rama de Bachillerato. Llamamos la atención al 20% de problemas descontextualizados donde el estudiante no puede ver el interés o la aplicación que pueda tener la estadística en su futura vida profesional.

- *Idoneidad ecológica.*

En este caso, se pretende medir el grado en que el proceso de estudio se ajusta al proyecto educativo de la institución y a los condicionamientos del entorno en que se desarrolla. Pensamos que esta idoneidad si la trata correctamente puesto que las PAU cumplen su función de evaluar los conocimientos y competencias de los estudiantes antes de su entrada a la universidad. Aunque sin duda el proceso se puede mejorar y existen propuestas de que cada facultad proponga sus propias pruebas este es un punto que necesita mucha reflexión y discusión, pues no todos los estudiantes en principio tienen claro los estudios que desean cursar en su nueva etapa universitaria.

En resumen, nuestro análisis indica que algunos componentes de la idoneidad didáctica son claramente mejorables, en particular, la idoneidad epistémica y cognitiva, por lo que debería ser tomada en cuenta por los diseñadores de las mismas en las sucesivas ediciones o en pruebas de evaluación alternativas que se propongan en el futuro.



# Capítulo 4

## CONCLUSIONES

Para finalizar la memoria, en este Capítulo se señalan, en base a los objetivos e hipótesis que sustentan el trabajo, los resultados más destacados. Asimismo, y teniendo en cuenta la investigación que se ha llevado a cabo, se realiza una breve reflexión sobre las limitaciones que presenta el trabajo y se proponen nuevas líneas de investigación con el fin de conseguir solventar las limitaciones encontradas. Considerando las características del trabajo, se finaliza el capítulo señalando las implicaciones que puede tener en la enseñanza y aprendizaje de los futuros estudiantes.

### 4.1 CONCLUSIONES RESPECTO A LOS OBJETIVOS DEL TRABAJO E HIPÓTESIS

Recordemos que en la Sección 1.5 del Capítulo 1 se señaló que este trabajo tiene como finalidad analizar los problemas de inferencia estadística propuestos en las PAU de la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II en la Comunidad de Andalucía en el periodo 2003-2014. Puesto que el objetivo general se desglosó en tres objetivos específicos, a continuación se realizará una breve reflexión sobre cada uno de ellos.

OE1. *Revisar las dificultades en la comprensión de los conceptos asociados a la inferencia estadística.* Dada la importancia que tiene el análisis de las dificultades que encuentran los estudiantes cuando se enfrentan a problemas de inferencia estadística, en el Capítulo 2 se ha realizado una revisión sobre las investigaciones llevadas a cabo que

más se relacionan con la nuestra. El estudio nos ha permitido comprobar la importancia que tiene este tema, ya que son numerosas las investigaciones encontradas, así como las dificultades que se describen. Entre otras señalamos la interpretación del valor  $p$ , del nivel de significación y de los resultados estadísticamente significativos, así como el planteamiento de las hipótesis adecuadas en un contraste de hipótesis.

El cálculo del intervalo de confianza es sencillo, pero su interpretación suele ser incorrecta. Por otro lado, los estudiantes confunden las diferentes distribuciones que intervienen en el proceso de inferencia: distribución en la población, distribución de los datos de la muestra y distribución muestral.

OE2. *Analizar el contenido sobre inferencia estadística propuesto en PAU.* Este objetivo se ha llevado a cabo mediante el análisis de las diferentes pruebas que han sido propuestas desde 2003 hasta 2014. En total se han resuelto y analizado 144 problemas, correspondiente a las 12 pruebas propuestas cada uno de estos años (6 en Junio y 6 en Septiembre). Por otro lado, teniendo en cuenta que cada problema tiene varios apartados a resolver, si se realiza el conteo según el número de apartados, este número asciende considerablemente hasta 285. La resolución de cada uno de los ejercicios y apartados y el análisis semiótico nos ha permitido analizar tanto los objetos matemáticos que intervienen en los problemas y compararlo con el contenido de las directrices curriculares sobre inferencia.

Como resultado se identificaron un gran número de objetos que intervienen en la resolución de los problemas propuestos (Véase Tabla 3.2), que prácticamente abarcan todos los contenidos curriculares de la inferencia en el periodo analizado (Tabla 3.3) y previstos en el nuevo Decreto (Tabla 3.4). No obstante no todos los contenidos se tienen en cuenta con la misma intensidad. Por un lado se privilegia el estudio de los intervalos de confianza y por otro la inferencia sobre la media de una población normal.

OE3. *Identificar las variables que definen los problemas de inferencia estadística y analizarlas estadísticamente.* El análisis empleado para examinar el contenido de inferencia nos ha permitido identificar las cuatro variables que han sido objeto de estudio. En primer lugar se identificaron cinco campos de problemas, analizando los objetos matemáticos requeridos en su solución y su relación con las directrices

curriculares. Además se analizó distribución de la población de partida, parámetro poblacional bajo estudio y contexto). Encontramos un gran énfasis en la media como parámetro y en la distribución normal; los contextos fueron variados, pero aun así se encuentran problemas no contextualizado lo cual contradice las recomendaciones actuales sobre la enseñanza de la estadística. El análisis llevado a cabo pone de manifiesto la importancia que tienen las variables identificadas en la elaboración de pruebas con características similares.

Con respecto a las hipótesis, que fueron consideradas como expectativas iniciales, se tienen las siguientes conclusiones:

*H1: El análisis de las investigaciones previas mostrará que el tema que nos ocupa ha sido escasamente investigado y, por tanto, nuestro trabajo puede aportar información original.* Aunque hemos encontrado bastantes trabajos que analizan las PAU desde diferentes puntos de vista, pocas se han centrado en el análisis de los problemas propuestos. En general estos trabajos analizan las características psicométricas, los resultados de los estudiantes o la opinión de los profesores.

La revisión llevada a cabo en el Capítulo 2 ha dejado al descubierto la escasez de investigaciones relacionadas con el análisis de las variables que determinan los problemas de inferencia estadística en las PAU. Los trabajos afines a la presente memoria han sido los presentados por Espinel, Ramos y Ramos (2006), Espinel, Ramos y Ramos (2007) y García y García (2005), sin embargo, destacamos que ninguno de ellos analiza las variables consideradas en esta investigación.

*H2: El análisis semiótico de los problemas de inferencia propuestos en las PAU y su comparación con el contenido curricular de estadística en los decretos curriculares de Bachillerato mostrará, por un lado, que se da una alta importancia a la inferencia en las PAU. Además se observará un buen ajuste entre los contenidos de inferencia fijados en el currículo y los considerados en las pruebas.* La inferencia estadística, al igual que la probabilidad, adquiere cierta importancia en las PAU ya que todas las pruebas incluyen un ejercicio que trata este tema. Por tanto determinan el 25% de la puntuación

en los estudiantes que realizan la prueba, a pesar de que el contenido de inferencia es menor que el 25% del currículo de matemáticas para estos estudiantes.

Además, el estudio llevado a cabo sobre los problemas de las PAU junto al análisis de las normativas curriculares pone de manifiesto, en general, una adecuada relación entre los contenidos tratados en las PAU e incluidos en el currículo. Sin embargo, hemos observado un mayor énfasis en la estimación de la media, en comparación con la estimación de la proporción y por otro lado los problemas propuestos se refieren a estimación en una sola población. Por tanto no se considera la comparación de dos poblaciones que es un tema incluido en el currículo.

*H3. Aunque en el currículo que han cursado los estudiantes en los pasados años (MEC, 2007) se sugiere estudiar la inferencia tanto para una como para dos muestras y en relación a la media y proporción, pensamos que la mayoría de problemas se centrarán tan sólo en la inferencia sobre la media de la muestra.* Como se ha expresado en la hipótesis anterior los resultados confirman nuestra hipótesis ya que no se han encontrado ejercicios donde se trabaje con dos poblaciones.

Puesto que el nuevo currículo (MECD, 2015) no incluye los contrastes de hipótesis una sugerencia sería eliminar los problemas relacionados con el mismo de las futuras pruebas de acceso a la universidad y completarlas con los temas escasamente tratados hasta ahora.

*H4. Al estudiar el contexto de los problemas, pensamos que se encontrarán una proporción grande de problemas descontextualizados.* El estudio estadístico confirma la expectativa inicial que se tenía. De los resultados hemos obtenido que el 21% de los ejercicios propuestos carecen de contexto. Sin embargo, es necesario notar que en los últimos años se está produciendo un cambio de tendencia con la inclusión de ejercicios enmarcados en un contexto social.

## **4.2. LIMITACIONES DEL ESTUDIO Y FUTURAS LÍNEAS DE INVESTIGACIÓN**

La principal limitación que se identifica de la investigación realizada es la dificultad que hay para poder acceder a las respuestas de los estudiantes ante estos problemas. Sería de gran interés analizar dichas respuestas para poder realizar una mejor clasificación de las dificultades que encuentran los estudiantes de Bachillerato cuando se enfrentan a un problema de inferencia y conseguir solventar dichas dificultades mediante una enseñanza adecuada.

Por otro lado, nos hemos limitado a los ejercicios propuestos en Andalucía. El estudio se podría ampliar comparando con los ejercicios de inferencia propuestos en las PAU de las distintas comunidades, lo que nos podría permitir realizar un análisis comparativo y examinar cómo se trata en todo el territorio español este tema.

Con el fin de contribuir a la manera cómo se forman los futuros universitarios en Estadística y Probabilidad, sería recomendable difundir los resultados aquí obtenidos con los recogidos en las investigaciones de Carretero (2014), Carretero, López-Martín y Arteaga (2015), Contreras, López-Martín, Arteaga y Carretero (2015), López-Martín, Contreras, Batanero y Carretero (2015) y López-Martín, Contreras, Carretero y Serrano (2016).

Igualmente se podría llevar a cabo un estudio de la presentación de la inferencia en los libros de texto de Bachillerato de Ciencias Sociales para comparar los ejercicios propuestos en los libros de texto y los utilizados en las PAU, y realizar recomendaciones para mejorar los textos, en caso de ser necesario.

## **4.3. IMPLICACIONES PARA LA ENSEÑANZA**

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 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.

Con el fin de garantizar el éxito de los estudiantes que se preparan para acceder a la universidad es vital asegurar una correcta y estrecha relación entre los contenidos de inferencia incluidos en las pruebas y los recogidos en los documentos curriculares. Los resultados obtenidos del análisis realizado en la investigación pueden servir para la elaboración de pruebas futuras (incluida la nueva reválida) y preparar a los estudiantes que tienen que enfrentarse a las mismas. Además, esperamos que la información recogida en esta memoria contribuya a mejorar la práctica docente de los profesores de la etapa de Bachillerato.

## REFERENCIAS

- Alvarado, H. y Batanero, C. (2007). Dificultades de comprensión de la aproximación normal a la distribución binomial. *Números*, 57. Online; <http://www.sinewton.org/numeros>.
- Arteaga, P. (2011). *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores*. Tesis Doctoral. Universidad de Granada.
- Batanero, C. (2000). Controversies around significance tests. *Mathematical Thinking and Learning*, 2(1-2), 75-98.
- Batanero, C. y Borovcnik, M. (2016). *Statistics and probability in high school*. Rotterdam: Sense publishers.
- Batanero, C. y Díaz, C. (2015). Aproximación informal al contraste de hipótesis. En J. M. Contreras (Ed.), *Actas de las II Jornadas Virtuales de Didáctica de la Estadística, la Probabilidad y la Combinatoria*. Granada: SEIEM. Online: <http://www.ugr.es/~batanero/documentos/Aproximacion.pdf>.
- Batanero, C. y Díaz, C. (2006). Methodological and didactical controversies around statistical inference: *Actes des 38-ièmes Journées de Statistique*. [CD-ROM] París: Société Française de Statistique.
- Batanero, C., Godino, J. D., Vallecillos, A., Green, D. R. y Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547.
- Batanero, C., Tauber, L. y Sánchez, B. (2001). Significado y comprensión de la distribución normal en un curso de análisis de datos. *Cuadrante*, 10(1), 59-92.
- Behar, R. (2001). *Aportaciones para la mejora del proceso de enseñanza-aprendizaje de la estadística*. Tesis Doctoral. Universidad Politécnica de Cataluña.
- Birnbaum, I. (1982). Interpreting statistical significance. *Teaching Statistics*, 4, 24-27.
- Bisquerra, R. (1989). *Métodos de investigación educativa*. Barcelona: CEAC.
- Blázquez, F. y Luengo, R. (1989). *Las calificaciones en las pruebas de acceso en la Universidad de Extremadura*. Badajoz: ICE, Universidad

- Boal, N., Bueno, C., Lerís, M. D. y Sein-Echaluce, M. L. (2008). Las habilidades matemáticas evaluadas en las Pruebas de Acceso a la Universidad. Un estudio en varias universidades públicas españolas. *Revista de Investigación Educativa*, 26(1), 11-23.
- Cabriá, S. (1994). *Filosofía de la estadística*. Valencia: Servicio de Publicaciones de la Universidad
- Caraballo, R. (2010). *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*. Trabajo fin de Máster. Universidad de Granada.
- Carretero, M. (2014). *Análisis de problemas de probabilidad en las pruebas de acceso de Matemáticas Aplicadas a las Ciencias Sociales de Andalucía*. Trabajo fin de Máster. Universidad de Granada.
- Carretero, M., López-Martín, M. M. y Arteaga, P. (2015). Contenido matemático de los problemas de probabilidad en las pruebas de acceso de Andalucía. *Probabilidad Condicionada. Actas de las II Jornadas Virtuales de Didáctica de la Estadística, la Probabilidad y la Combinatoria*. Granada: SEIEM. Online: <http://www.jvdiesproyco.es/>.
- Castro-Sotos, A. E., Vanhoof, S., Noortgate, W. y Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review*, 2(2), 98-113.
- Castellanos, M. T. (2013). *Tablas y gráficos estadísticos en las pruebas Saber Colombia*. Trabajo fin de Máster. Universidad de Granada.
- Chance, B., delMas, R. C. y Garfield, J. (2004). Reasoning about sampling distributions. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295-323). Amsterdam: Kluwer.
- Colera, J., García, R. y Oliveira, M. J. (2003). *Matemáticas aplicadas a las Ciencias Sociales II*. Madrid: Anaya.
- CECD, Consejería de Educación, Cultura y Deporte de la Junta de Andalucía (2015). *Decreto 97/2015, de 3 de marzo, por el que se establece la ordenación y el currículo de la Educación Primaria en la Comunidad Autónoma de Andalucía*. Sevilla: Autor.

- Contreras, J. M., López-Martín, M. M., Arteaga, P. y Carretero, M. (2015, Febrero). Probability content in the entrance to university tests in Andalusia. Presentado en International Conference. *Turning data into knowledge: new opportunities for Statistics Education*. Lisboa: Instituto de Educaçao.
- Contreras, A., Ordoñez, L. y Wilhelmi, M. R. (2010). Influencia de las Pruebas de Acceso a la Universidad en la enseñanza de la integral definida en el Bachillerato. *Enseñanza de las Ciencias* 28(3), 367-384.
- 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.
- Cumming, G. y Fidler, F. (2005, Agosto). Interval estimates for statistical communication: problems and possible solutions. Trabajo presentado en la *IASE Satellite Conference on Communication of Statistics*. Sydney: International Association for Statistical Education.
- Cuxart, A. (2000). Modelos estadísticos y evaluación: tres estudios en educación. *Revista de Educación*, 323, 369-394
- Díaz, C., Contreras, J. M. Batanero, C. y Roa, R. (2012). Evaluación de sesgos en el razonamiento sobre probabilidad condicional en futuros profesores de educación Secundaria. *Bolema* 26 (22), 1207-1226.
- Díaz, C. y de la Fuente, I. (2005). Razonamiento sobre probabilidad condicional e implicaciones para la enseñanza de la estadística. *Epsilon*, 59, 245-260.
- Díaz, P., Mier, V., Alonso P. y Rodríguez-Muñiz; L. J. (2014). Probability and statistics in access exams to Spanish universities. En K. Makar, B. de Sousa y R. Gould (Eds.), Sustainability in statistics education. *Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014)*, Flagstaff, AR: International Statistical Institute.
- Espinel, M. C., Ramos, C. E. y Ramos. R. M. (2006, Septiembre). La inferencia estadística en la PAU. Presentado en *XXIX Congreso Nacional de Estadística e Investigación Operativa, SEIO*. Tenerife.
- Espinel, M. C., Ramos, R. M. y Ramos, C. E. (2007). Algunas alternativas para la mejora de la enseñanza de la inferencia estadística en Secundaria. *Números*, 67, 15-23.

- Falk, R. (1986). Conditional probabilities: insights and difficulties. En R. Davidson y J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics*. (pp. 292 – 297). Victoria, Canada: International Statistical Institute.
- Falk, R. y Greenbaum, C. W. (1995) Significance tests die hard: The amazing persistence of a probabilistic misconception, *Theory and Psychology*, 5(1), 75-98.
- Fidler, F. (2005). *From statistical significance to effect estimation*. Tesis Doctoral. Universidad de Melbourne.
- Fidler, F. y Cumming, G. (2005). Teaching confidence intervals: Problems and potential solutions. *Proceedings of the 55th Session of the International Statistical Institute* (pp. 1-5). Voorburg: International Statistical Institute. Online: [iase-web.org/documents/papers/isi55/Fidler-Cumming.pdf](http://iase-web.org/documents/papers/isi55/Fidler-Cumming.pdf).
- Fisher, R. A. (1935). *The design of experiments*. Edimburgh: Oliver y Boyd.
- Font, V., Godino, J. D., y Gallardo, J. (2013). The emergence of objects from mathematical practices. *Educational Studies in Mathematics*, 82, 97–124.
- Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M. y Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K-12 curriculum framework*. Alexandria, VA: American Statistical Association. Online: [www.amstat.org/Education/gaise/](http://www.amstat.org/Education/gaise/).
- García, I. y García, J.A. (2005). Algunos resultados sobre la actuación de los alumnos en las cuestiones de estadística en la P. A. U. *Actas de las XI Jornadas para el Aprendizaje y Enseñanza de las Matemáticas* (pp. 733-738). La Laguna: Consejería de Educación, Cultura de Deportes.
- Garfield, J. y Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. New York: Springer Science & Business Media.
- Gaviria, J. L. (2005). La equiparación del expediente de Bachillerato en el proceso de selección de alumnos para el acceso a la universidad. *Revista de Educación*, 337, 351-387.
- Goetz, J. P. y Lecompte, M. D. (1998). *Etnografía y diseño cualitativo en educación*. Morata: Madrid.
- 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 Mathematiques*, 22(2-3), 237-284.
- Godino, J. D. (2011, Julio). Indicadores de la idoneidad didáctica de procesos de enseñanza y aprendizaje de las matemáticas. Trabajo presentado en la *XIII CIAEM-IACME*, Recife, Brasil.
- 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.
- Godino, J.D., Bencomo, D., Font, V., y Wilhelmi, M.R. (2006) Análisis y valoración de la idoneidad didáctica de procesos de estudio de las matemáticas. *Paradigma*, 27(2), 221-252.
- Godino, J. D., Contreras, Á. y Font, V. (2006). Análisis de procesos de instrucción basado en el enfoque ontológico-semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, 26(76), 39.
- Grau, R., Cuxart, A. y Martí-Recober, M. (2002). La calidad en el proceso de corrección de las Pruebas de Acceso a la Universidad: variabilidad y factores. *Revista de Investigación Educativa*, 20(1), 209-223.
- Hald, A. (2008). *A history of parametric statistical inference from Bernoulli to Fisher, 1713-1935*. New York: Springer.
- 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.
- Inzunsa, S. y Jiménez, J. V. (2013). Caracterización del razonamiento estadístico de estudiantes universitarios acerca de las pruebas de hipótesis. *Revista Latinoamericana de Investigación en Matemática Educativa*, 16(2), 179-211.
- Krauss, S. y Wassner, C. (2002). How significance tests should be presented to avoid the typical misinterpretations. En C. Reading (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town, South Africa: International Association for Statistics Education. Online: [www.stat.auckland.ac.nz/~iase/publications](http://www.stat.auckland.ac.nz/~iase/publications).
- Krippendorff, K. (2013). *Content analysis: an introduction to its methodology*. London, Sage

- Lecoutre, B. y Poitevineau, J. (2014). *The significance test controversy revisited*. Berlin: Springer.
- Liu, Y. y Thompson, P. W. (2009). Mathematics teachers' understandings of proto-hypothesis testing. *Pedagogies*, 4 (2), 126-138.
- López-Martín, M. M., Contreras, J. M., Batanero, C. y Carretero, M. (2015). Los problemas de probabilidad propuestos en las Pruebas de Acceso a la Universidad en Andalucía. *Areté*, 1(1), 39-60.
- López-Martín, M. M., Contreras, J. M., Carretero, M., Serrano, L. (2016). Análisis de los problemas de probabilidad propuestos en las Pruebas de Acceso a la Universidad en Andalucía. *Avances de Investigación en Educación Matemática (AIEM)*, 9, 65-84.
- Mallart, A. (2014). La resolución de problemas en la prueba de Matemáticas de acceso a la universidad: procesos y errores. *Educatio Siglo XXI*, 32(1), 233-254.
- Mengual, E., Gorgorió, N. y Albarracín, L. (2013). Validación de un instrumento para la calificación de exámenes de matemáticas. *Investigación en Educación Matemática XVI* 367-381.
- Mingorance, C. (2014). *La estadística en las pruebas de diagnóstico andaluzas*. Trabajo fin de Grado. Universidad de Granada.
- MCD, Ministerio de Educación, Cultura y Deporte (2013). *PISA 2012. Programa para la evaluación internacional de los estudiantes. Informe español. Vol.1. Resultados y contexto*. Madrid: Autor.
- ME, Ministerio de Educación (2009). *PISA 2009. Programa para la evaluación internacional de estudiantes de la OCDE. Informe español*. Madrid: Autor.
- 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 (1999). *Real Decreto 1640/1999, de 22 de octubre, por el que se regula la prueba de acceso a estudios universitarios*. Madrid: Autor.
- MECD, Ministerio de Educación, Cultura y Deporte (2000). *Real Decreto 990/2000, de 2 de junio, por el que se modifica y completa el Real Decreto 1640/1999, de 22 de octubre, por el que se regula la prueba de acceso a estudios universitarios*. Madrid: Autor.

- MECD, Ministerio de Educación, Cultura y Deporte (2002). *Real Decreto 1025/2002, de 4 de octubre, por el que se modifica el Real Decreto 1640/1999, de 22 de octubre, modificado y completado por el Real Decreto 990/2000, de 2 de junio, por el que se regula la prueba de acceso a estudios universitarios*. Madrid: Autor.
- MECD, Ministerio de Educación, Cultura y Deporte (2003). *Real Decreto 1741/2003, de 19 de diciembre, por el que se regula la prueba general de Bachillerato*. Madrid: Autor.
- MECD, Ministerio de Educación, Cultura y Deporte (2014). *Real Decreto 412/2014, de 6 de junio, por el que se establece la normativa básica de los procedimientos de admisión a las enseñanzas universitarias oficiales de Grado*. Madrid: Autor.
- MECD, Ministerio de Educación, Cultura y Deporte (2015a). *Proyecto de real decreto por el que se regulan las características generales de las pruebas de la evaluación final de Educación Primaria, y las características de las pruebas de las evaluaciones finales de Educación Secundaria Obligatoria y de Bachillerato, establecidas en la Ley Orgánica 2/2006, de 3 de mayo, de Educación*. Madrid: Autor.
- MECD, Ministerio de Educación, Cultura y Deporte (2015b). *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.
- Moore, D. S. (2010). *The basic practice of statistics*. New York: Freeman (5th edition).
- 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.
- Muñoz Vitoria, F. (1995). El acceso a la Universidad en España: perspectiva histórica. *Revista de Educación*, (308), 31-61.
- Murillo, F. J. (1997). Análisis de las pruebas que conforman la selectividad. *Revista de Educación*, 314, 49-62.
- Neyman J. y Pearson E. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society A*, 231, 289–337

- Nortes, A. N. y Nortes, R. (2010). Resolución de problemas de matemáticas en las Pruebas de Acceso a la Universidad. Errores significativos. *Educatio Siglo XXI*, 28(1), 317-342.
- Nortes, A. N., Nortes, R. y Lozano, F. (2015). Las correcciones en Matemáticas en las Pruebas de Acceso a la Universidad. *Educatio Siglo XXI*, 33(3), 199-222.
- OCDE, Organización para la Cooperación y el Desarrollo Económico (2015). *PISA 2015 mathematics framework*. Paris: OCDE
- Popper, K. (1959). *The logic of scientific discovery*. London: Hutchinson.
- Ramos, C. E., Espinel, M.C. y Ramos. R. M. (2009). Identificación de los errores en los contrastes de hipótesis de los alumnos de Bachillerato. *SUMA*, 61, 35-44.
- Rico, L. (2006a). Marco teórico de evaluación en PISA sobre matemáticas y resolución de problemas. *Revista de Educación*, (Extraordinario 2006), 275-294.
- Rico, L. (2006b). La competencia matemática en PISA. *PNA*, 1(2), 47-66.
- Rossmann, A. (2008). Reasoning about informal statistical inference: One statistician's view. *Statistics Education Research Journal*, 7 (2), 5-19. Online: [iase-web.org/Publications.php?p=SERJ](http://iase-web.org/Publications.php?p=SERJ).
- Rivadulla, A. (1991). *Probabilidad e inferencia científica*. Barcelona: Anthropos.
- Ruiz, J., Dávila, P., Etxeberria, 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.
- Ruiz, J., Sarasua, J. y García, J. M. (2011). Una tipología y clasificación de los ejercicios de matemáticas de selectividad. *Epsilon*, 78, 21-38.
- Saldanha. L. y Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257-270.
- Saldanha. L. y Thompson, P. (2007) Exploring connections between sampling distributions and statistical inference: an analysis of students' engagement and thinking in the context of instruction involving repeated sampling. *International Electronic Journal of Mathematics Education*, 2(3), 270-297. Online: [www.iejme.com/](http://www.iejme.com/).
- 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.
- Tukey, J.W. (1962). The future of data analysis. *The Annals of Mathematical*

*Statistics*, 33(1), 1-67.

Vallecillos, A. (1994). *Estudio teórico experimental de errores y concepciones sobre el contraste de hipótesis en estudiantes universitarios*. Tesis doctoral. Universidad de Granada,

Vallecillos, A. (1999). Some empirical evidence on learning difficulties about testing hypotheses. *Proceedings of the 52 session of the International Statistical Institute* (Vol.2, pp. 201–204). Helsinki: International Statistical Institute.

Weber, R. P. (19). *Basic content analysis*. Londres: Sage.